PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

FEB 24 1930

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

HAROLD RUGG, *Editor*

VOL. XXI
No. 2

FEBRUARY, 1930

$5.00 a year.
70c per copy.

CONTENTS

HAPPINESS AMONG ADULT STUDENTS OF EDUCATION. *Goodwin Watson* 79

A TRUCE TO "BARKING IN." *C. Spearman* 110

FURTHER STUDIES OF CHILDREN'S INTEREST IN COLLECTING. *Paul A. Witty and Harvey C. Lehman* 112

AN IMPROVEMENT IN THE THEORETICAL BASIS OF FIVE POINT GRADING SYSTEMS BASED ON THE NORMAL PROBABILITY CURVE. *Walter Crosby Eells* 128

WEIGHTING RESPONSES IN TRUE-FALSE EXAMINATIONS. *Elmer H. Staffelbach* 136

A GRAPH FOR ESTIMATING RELIABILITY IN ONE RANGE, KNOWING IT IN ANOTHER. *Phillip Justin Rulon* 140

THE INFLUENCE OF WEIGHTED TRUE-FALSE TEST SCORES ON GRADES. *John Gray Peatman* 143

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 148

PUBLICATIONS RECEIVED 156

Published Monthly Except June to August by

Warwick & York, Inc.,

10 East Centre Street Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at Baltimore, Md., under the Act of March 3, 1879; additional entry as second class matter at York, Pa.

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

ESTABLISHED 1910

Editor: HAROLD RUGG

Board of Associate Editors

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — L. M. TERMAN
FRANK N. FREEMAN — V. A. C. HENMON — GOODWIN B. WATSON — E. L. THORNDIKE

Assistant to the Editor: FRANCES M. FOSTER

The Journal is published monthly, except June to August, by Warwick and York, Inc., Baltimore, Md., and York, Pa.

The price is $5.00 a year, payable in advance; foreign postage 30 cents extra. Single current issues are 70 cents each, and subscriptions for less than a year are charged at single issue rate for each number ordered. Back volumes are $6.00 each; single back numbers, 90 cents each, except, when more than three years old, the price is $1.00 each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of the next succeeding number.

Manuscripts for publication, books or other materials for review, and correspondence concerning editorial matters should be addressed to Harold Rugg, Editor, 425 West 123rd Street, New York City. Subscriptions should be sent to the publishers.

WARWICK AND YORK, INC., *Baltimore, Maryland*

INTRODUCTORY PSYCHOLOGY FOR TEACHERS

By EDWARD K. STRONG, JR.

$2.20 + 12¢ postage

WARWICK & YORK, INC., *Baltimore, Maryland*

TEACHERS' HEADQUARTERS

European Plan

Entirely Fire-Proof

Centrally Located

EDWARD DAVIS

Manager

HOTEL RENNERT

BALTIMORE, MARYLAND

SCHOOL AND SOCIETY

A weekly journal covering the field of education in relation to the problems of American democracy.

The objects of SCHOOL AND SOCIETY are the advancement of education as a science and the adjustment of our lower and higher schools to the needs of modern life.

It encourages discussion of all problems from every point of view; it pays special attention to foreign affairs; it publishes accounts of educational events of the past week and research on fundamental problems.

Subscription $5.00—Single Numbers 15 Cents

SCHOOL AND SOCIETY will be sent to new subscribers for thirteen weeks for one dollar, which may be sent at the risk of the publisher.

THE SCIENCE PRESS

Grand Central Terminal - - New York, N. Y.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXI February, 1930 Number 2

HAPPINESS AMONG ADULT STUDENTS OF EDUCATION

GOODWIN WATSON

Associate Professor of Education, Teachers College, Columbia University

No human quest may claim a larger following than that for happiness and satisfaction in life. Even the highest ethics tends to justify itself by its contribution to human happiness. The Beatitudes themselves constitute an attempt to formulate the conditions of the most blessed existence. Certainly any educational program in modern times is likely to be justified only in terms of its direct or indirect contribution to human happiness. It becomes, therefore, extraordinary almost beyond belief that so few attempts have been made to apply the techniques of psychological study to the understanding of happiness.

This study is based upon the responses of 388 graduate students of education, averaging thirty years of age. What is studied might, in the strictest sense be termed not happiness, but self-estimates of happiness. There is a sense in which an individual may be said to err in his judgment of his own happiness. Perhaps he doesn't realize how "happy" he is. Probably "fortunate" would be a better word for that connotation, however.

In this study, the individual is called happy if he believes himself happier than most others of like age and sex, if he believes his prevailing moods cheerful, his spirits high, his satisfactions lasting, his days full of interesting and amusing things, his prevalent attitudes described by such words as "enthusiastic," "jolly," "tranquil," "joyful," "fortunate," or "well-integrated." If he reports splendid satisfaction with his health, his work, his love adjustment, his friends, hobbies, and religion, these likewise contribute to his "happiness" score. Individuals are called unhappy, whatever their status in the sight of Omniscience, if they believe themselves generally unhappy, describe their prevalent moods

as disillusioned and depressed, life seeming so worthless that there is little to keep them going, are usually sad, weep readily, smile seldom, are morbid, disappointed, melancholy, blue, lonely, troubled, miserable, confused, feel that nothing matters or that they have been seriously hurt by life so that the good aspects of living never really compensate for the injuries. If health is poor, vocation unsatisfying or a failure, love adjustment unsatisfactory, friends a source of bitterness, hobbies lacking or annoying, religion a source of conflict and worry, these factors were taken into account as contributory to general unhappiness. Most students, of course, fall between these two extremes.

The report of happiness was obtained by the use of a blank, part of which is reproduced herewith.[1] Copies were distributed, the purpose of the study explained, and practically complete returns secured. A card attached to the questionnaire carried the name of the individual, and the number of his blank. These were turned in separately, kept very confidential, and used only to ascertain intelligence scores for each individual. All cards were then destroyed. The understanding was that the study was anonymous so far as any individual was concerned.

An Estimate of Happiness

1. (*a*) Comparing yourself with other persons of the same age and sex how do you feel you should rate your own general happiness? Place a short vertical mark across the line below to indicate about where you belong. Consider your average state over several months.

Most miserable of all	About three-fourths of the population happier than you are	The average person of your own age and sex	Happier, on the whole, than three-fourths of the population of similar age and sex	Happiest of all

(*b*) Place a circle on the line to indicate about where you believe your friends would rate you, whether at the same point at which you rate yourself or higher or lower.

2. Among the following descriptions, arranged in miscellaneous order, choose the one which comes nearest to fitting you. None will be likely to fit exactly. Place here the letter of the one which seems to you most appropriate. [ ]

(*a*) Finding life rather disappointing and disillusioning, comfortable in many ways, moderately successful, but far from realizing the hopes of youth.

---

[1] Copies of the complete form may be secured through the author.

(*b*) Cheerful, gay spirits most of the time. Occasionally bothered by something but can usually laugh it off.

(*c*) Calm, quiet sort of satisfaction. Life has been pretty good. Not everything one desires comes, of course, but on the whole there is much for which to be serenely thankful.

(*d*) Ups and downs, now happy about things, now depressed. About balanced in the long run.

(*e*) Life often seems so worthless that there is little to keep one going. Nothing matters very much, there has been so much of hurt that laughter would be empty mockery.

(*f*) Keeping a brave front, others think everything is all right. Inside life seems rather black.

(*g*) Usually sad, weep readily, smile seldom.

(*h*) Radiant, find every day full of interest, amusing things, and worthwhile things.

(*i*) Seriously hurt by certain things, for which the good aspects of life can not quite make up.

(*j*) Quite objective. Like some experiences, dislike others. Not aware of any prevalent happiness or unhappiness.

3. Now write in your own words a sentence or two, something like those above, which you believe will most truly describe your own general happiness in life.

.....

.....

.....

.....

.....

4. Below are a list of words and phrases. Check every term which you believe could fairly be applied to yourself in your prevalent attitudes.

——Enthusiastic	——Morbid	——Disappointed
——Distressed	——Cheerful	——Prosperous
——Frivolous	——Troubled	——Annoyed
——Calm	——Miserable	——Thrilled
——Irritable	——Buoyant	——Joyful

(The total list contained fifty words half positive half negative.)

5. Consider now the various aspects of your living mentioned below. Rate your happiness in each one, in comparison with the general population of the same age and sex. Place a short vertical mark across the line to indicate about where you believe you belong.

(Here followed graphic rating scales similar to the one used in the first question. These covered the areas of health, vocation, love and marriage, friends, hobby interests and religion.)

6. You are now asked to rate your happiness by various stages of your life. For each of the age periods below, indicate your judgment of your enjoyment of life in comparison with that of other children of the same age and sex.

(Graphic rating scales, similar to those already used in the study, covered early childhood (age three to seven inclusive), later childhood (eight to twelve inclusive), high school period, and later adolescence.)

The happiness score was computed by assigning values to each question in an arbitrary fashion, as a result of discussion by a committee who had examined and tabulated the responses.[1] Each graphic rating scale was scored by a scale of units ranging from 0 at the most unhappy extreme to 100 at the happiest extreme with 50 in the middle. The descriptions in question 2 were assigned values, respectively, of 3, 9, 8, 5, 1, 2, 1, 10, 3, and 5. This was then weighted, being multiplied by 7. The subjective statement was scored by three judges, the median rating being used. Judgments ranged from 1 to 10. The resulting judgment was multiplied by 6 to give it a proper weighting. The score on the adjectives was obtained by subtracting the number of "unhappy" traits mentioned from the number of "happy" ones. Happiness in the six areas of experience was included by averaging the graphic score for the six items, and using only the average in making up the happiness score. Table I indicates the relative weight of each factor in making up the composite. Most weight was assigned to the average of six ratings in special areas of experience, next most to the choice of suggested descriptions and the general rating. Least was

TABLE I.—INFLUENCE OF EACH QUESTION UPON TOTAL HAPPINESS SCORE, AS INDICATED BY ITS VARIABILITY

FACTOR	SD MEN (N = 50)	SD WOMEN (N = 124)
1. Self-estimate of general happiness	15.3	13.2
2. Choice of suggested description	15.1	17.2
3. Original description of prevalent mood	9.6	10.8
4. Adjectives	6.2	6.9
5. Happiness in each of special areas (averaged and used only once)		
Health	15.7	13.8
Vocation	17.0	14.3
Love	20.5	19.4
Friends	15.5	13.2
Hobbies	16.1	15.7
Religion	19.5	18.3
6. Happiness at various stages of life (averaged and used only once)		
Early childhood	17.5	14.8
Later childhood	18.2	17.6
High school period	16.3	18.0
Later adolescence	20.3	18.5

[1] The author wishes to express his gratitude to the committee in the Psychology of Character course which assisted in the tabulation and interpretation of this study: Richard Hull, Marion Streng, Belle Holaway, L. Earl Willmott, Mary Willmott, and Helen Hazleton.

assigned to the original description which had to be somewhat subjectively scored, and to the adjectives, choice of which appeared to be somewhat unreliable.

Incidentally, Table I also suggests that there is little sex difference in variability upon such indices. In all of the graphic ratings the men were slightly more variable than the women; in the choice of descriptions, the description in the subject's own words, and in the checking of adjectives the women were more variable. Love and religion appeared to be the areas calling forth greatest diversity in answer between the happy and the unhappy for both sexes.

Reliability of this total score was studied by obtaining a score (*a*) made up of general estimate of happiness, free description, rating on happiness in health, love and hobbies, and another score (*b*) made up of choice of suggested descriptions, checking of adjectives, and estimate of happiness in vocation, friendship, and religion. Correlation of these two scores from seventy-five men and one hundred twenty-five women with correction by the Spearman-Brown formula gave reliabilities, respectively of .83 and .85 This indicates that the measure is reasonably consistent, throughout. It does not, of course, answer the other very interesting but as yet unstudied question of the stability of this measure from day to day, and from week to week.

TABLE II.—DISTRIBUTION OF HAPPINESS SCORES

SCORES	FREQUENCY
361–380	7
341–360	17
321–340	33
301–320	53
281–300	63
261–280	65
241–260	34
221–240	32
201–220	39
181–200	18
161–180	15
141–160	8
121–140	2
100–120	2
	$N = 388$

Q_1, 226. Median, 273. Q_3, 305. $Q = 40$.

The resulting happiness scores are shown in the distribution, Table II. This may take on significance in terms of the "theoretically average" person. One who regarded his happiness in general and in each area of life as average, who described himself by saying, "ups and

downs, now happy, now depressed, about balanced in the long run," or "quite objective; like some experiences, dislike others, not aware of any prevalent happiness or unhappiness," checking no more optimistic adjectives to describe himself than he checked troubled adjectives, would earn a happiness score of 215. Eighty-one per cent of this group of graduate students of education surpassed this score. In general, they tended to think of themselves as above the average on each happiness item. The distribution itself is not far from normal, although skewed slightly in the direction of the higher scores.

The median graduate student might, perhaps, have attained his score by such a self-rating as this. He might regard himself, as surpassing about three-fourths of his age and sex in general happiness, might check as most nearly describing himself, "calm, quiet sort of satisfaction; life has been pretty good; not everything one desires, of course, but on the whole there is much for which to be serenely thankful." He might himself write, as did one individual (score 274) by saying, "Happy in work. Often a little discouraged at not accomplishing what I would like to do most. Then at times highly pleased at certain results which I have worked for." On the adjective list, such an average student might check calm, irritable, well-balanced, good sense of humor, glad, tranquil, melancholy, pleasant, successful, annoyed, peaceful, and fortunate. Health would be regarded as good, vocation practically ideal; love and marriage, below average; friends highly satisfactory, hobbies enjoyable; religion, entirely satisfactory. Happiness in early childhood might be rated high, in later childhood average, in adolescence low, in recent years back again above average. Such is the average level of happiness portrayed by these adults.

A study of the relationship between scores on various parts of this index, and total happiness score may serve two purposes. It may indicate ways in which such a self-estimating device can be abbreviated. If all indices measure essentially the same thing, the first device may be needlessly complicated. Again, such a comparison may show something about the factors which are most important for happiness. This is particularly true when happiness in each of the several areas is compared with happiness in general.

Table III makes this comparison by selecting three groups, the fifty happiest, fifty near the median, and the fifty most unhappy. The mean score of each group on each index is given. Thus it appears that the check mark on the first graphic scale, indicating general happiness

was placed at 35 by the unhappy group, 53 by the average group, and 69 by the happiest group. This factor is said to be most clearly related to whatever is measured by the total happiness score, because it separates the happy and unhappy groups by a larger difference than does any other factor. Love and marriage stands second, separating

TABLE III.—SEVERAL ESTIMATES OF HAPPINESS ARRANGED IN ORDER OF DIFFERENTIATING POWER[1]

	Fifty most unhappy	Fifty nearest median	Fifty happiest	Difference between happy and unhappy
1. Estimate of present general happiness; average state over several months	35	53	69	34
2. Estimate of happiness in love and marriage; how far ideal	34	64	65	31
3. Estimate of happiness in vocation; how far ideal for subject	45	71	70	25
4. Indication of how subject would expect to be rated on happiness by his friends	46	62	69	23
5. Estimate of happiness in hobby interests, how enjoyable	52	62	74	22
6. Estimate of happiness in religion, how satisfying	46	57	67	21
7. Estimate of general happiness during later adolescence, age 18–25	44	52	63	19
8. Estimate of happiness in health, how far free from illness, energetic, sense of euphoria	58	65	75	17
9. Estimate of happiness and satisfaction in relations with friends, how nearly ideal	51	61	66	17
10. Estimate of general happiness in later childhood, 8–12 years	44	52	56	12
11. Estimate of general happiness in early childhood, 3–7 years	52	55	59	7
12. Estimate of general happiness during high school age, 13–17 years	49	50	53	4

[1] All were indicated by marks on a scale in which 0 represented the most unhappy person, 50 an average person of like age and sex, 100 the happiest such person. The SD difference in any comparison lies between 2 and 3. Differences of 7 and over are usually dependable.

the unhappy group (34) from the happy group (65) by a large margin, but failing to differentiate between the average group (64) and the happiest group (65). Happiness in vocation is likewise low for the generally unhappy group, but almost identical in the median and happy groups. Estimates of happiness in early childhood and during the high school period tend to be almost unrelated to present happiness score.

TABLE IV.—RELATIONSHIP OF ASPECTS OF HAPPINESS TO TOTAL HAPPINESS SCORE

	MEN (N = 50)	WOMEN (N = 124)
1. Self-rating on general happiness, average state over several months	.813	.818
2. Choice of suggested description of mood and temperament	.686	.830
3. Subject's own description of his general happiness in life	.773	.735
4. Adjectives applicable to subject's prevalent attitudes	.657	.782
5. Self-rating on satisfaction with vocation	.714	.511
6. Estimate of how happiness would be rated by friend	.466	.560
7. Self-rating on satisfaction with religion	.414	.485
8. Self-rating on satisfaction with love and marriage	.616	.368
9. Self-rating on satisfaction with friends	.457	.390
10. Self-rating on satisfaction with health	.276	.477
11. Self-rating on satisfaction with hobby interests	.062	.305

TABLE V.—RELATIONSHIP OF PRESENT HAPPINESS TO HAPPINESS AT VARIOUS PREVIOUS PERIODS

	MEN (N = 50)	WOMEN (N = 124)
1. Self-rating of happiness in early childhood, age 3–7	.133	.080
2. Self-rating of happiness in later childhood, age 8–12	.127	.224
3. Self-rating of happiness in high school period	.315	.154
4. Self-rating of happiness in later adolescence, 18–25	.330	.278

Tables IV and V study the same problem using correlations for the entire distribution. It appears that the single check mark on the graphic rating scale would give a correlation of .81 or .82 with this entire happiness scoring. For the purpose of separating high and low groups such a simple indicator would probably be adequate. The choice of suggested descriptions stands next in general efficiency, but apparently the mood descriptions were better adapted to distinguishing the women than the men. The subject's own description, although scored subjectively, would appear to yield a correlation over .7 with the total happiness. Of course, all of these correlations may be said to be spuriously high because each is included in the total. Among the areas stated specifically, vocational satisfaction (.71 and .51) showed the highest relationship, religion (.41 and .49) and love (.62 and .37)

running a poor second and third. Hobby interests and health would appear to be little related to the total. Table V indicates that the matter of happiness at other periods of life could well be dropped from the schedule. Memory may be inaccurate, it may be compensatory, or there may be no relationship between present and earlier happiness. In any case there appears to be little agreement between estimates of happiness during the childhood and high school periods, and present general happiness.

Sex differences are not clearly distinguishable in Tables IV and V because of the small number of cases. The standard error of the difference is such that only differences of .2 or more are clearly beyond the possibility of a reverse direction with another sampling. It does seem reasonably certain that health and hobby interests were more clearly related to happiness among women than among men, that vocational satisfaction and satisfaction in love and marriage were more clearly related to general happiness among men than among women.

In addition to estimating his general happiness, in the first question, each subject was asked to indicate how his friends would rate him. It is of passing interest to examine the results, first by the method of group comparisons, then by the correlation technique.

TABLE VI.—AVERAGE SELF-ESTIMATES AND ESTIMATES ATTRIBUTED TO FRIENDS

	Fifty Most Unhappy by Total Score	Fifty Nearest Median in Total Score	Fifty Happiest by Total Score
Check indicating their own estimate of their own happiness	35	53	69
Circle indicating how they believe their friends would rate them	46	62	69

Both the low and the average groups appeared to believe that they gave an impression of greater happiness than was justified. It would be interesting to check this assumption against actual ratings by friends. At any rate, the happy group felt no sense of masquerade. In terms of correlation, Table VII shows that there was a relationship amounting to .47 or .56 between total happiness score and the scores which these students believed would be assigned them by their friends. Little relationship appeared between this estimate of friends' ratings and any of the other factors with which it is there compared.

Scores on intelligence tests are so widely used in educational studies and so often popularly misunderstood to be equivalent to chance for success and satisfaction in life, that it seemed wise to make a

TABLE VII.—CORRELATIONS WITH BELIEF ABOUT HOW FRIENDS WOULD RATE SUBJECT'S HAPPINESS

CORRELATION WITH	MEN (N = 50)	WOMEN (N = 124)
1. Total happiness score	.446	.560
2. Subject's self-estimate of average happiness	.528	.554
3. Choice of suggested description of mood and temperament	.261	.491
4. Subject's own free description of his general happiness in life.	.222	.348
5. Adjectives applicable to subject's prevalent attitudes	.386	.388
6. Self-rating on satisfaction with health	.085	.226
7. Self-rating on satisfaction with vocation	.281	.320
8. Self-rating on satisfaction with love and marriage	.462	.056
9. Self-rating on satisfaction with friends	.216	.345
10. Self-rating on satisfaction with hobby interests	.017	.148
11. Self-rating on satisfaction with religion	.299	.201
12. Self-rating of happiness in early childhood: age 3–7	.131	.048
13. Self-rating of happiness in later childhood: age 8–12	.090	.033
14. Self-rating of happiness in high school period	.082	.038
15. Self-rating of happiness in later adolescence 18–25	.030	.104
16. Otis Intelligence Test score	−.111	.017
17. Self-estimate of popularity with same sex in high school	.244	.091
18. Self-estimate of popularity with opposite sex in high school	.304	−.009

special study of this point. Is it true, as some have assumed, that the dull have little chance at happiness? Or, is it true, on the converse, that the moron is the happy creature? These data do not cover a wide enough sampling of the population by intelligence, age, or occupation to permit any general answer. For this group of teachers, however, Tables VIII and IX give a clear cut answer. Intelligence has nothing at all to do with the matter. Table VIII shows approximately the same happiness range and midscore for each tenth of the total group in intelligence. Those in the brightest groups ranged from

TABLE VIII.—MEDIAN HAPPINESS SCORE FOR EACH DECILE GROUP IN OTIS INTELLIGENCE TEST

OTIS TEST	MIDSCORES ON HAPPINESS	RANGE IN HAPPINESS SCORE
Highest decile	278	112–350
Ninth decile	275	117–355
Eighth decile	274	147–339
Seventh decile	276	141–343
Sixth decile	294	173–364
Fifth decile	267	166–355
Fourth decile	285	184–359
Third decile	274	154–345
Second decile	294	167–352
Lowest decile	276	147–346

less than 120 to over 340. Those in the dullest groups range from less than 170 to over 340. Table IX shows a correlation of −.032 in the case of fifty men, a correlation of −.062 in the case of one hundred twenty-four women, for the relationship between score on the Otis S-A Test of Mental Ability, Higher Examination, and total happiness score as obtained in this study. Table IX shows likewise that intelligence test score is unrelated to each of the sub-indices of happiness. It did not influence choice of description of prevailing mood, choice of adjectives, satisfaction in realms of health, or vocation, or love, or friends, or hobbies, or religion. The last two items in the table show intelligence to be similarly unrelated to the subject's estimate of his popularity in high school, either with the same sex or with the opposite sex.

TABLE IX.—CORRELATION WITH INTELLIGENCE (SCORE ON OTIS S-A TEST OF MENTAL ABILITY, HIGHER EXAMINATION)

CORRELATED WITH	MEN (N = 50)	WOMEN (N = 124)
1. Total happiness score	−.032	−.062
2. Subject's self-estimate of general happiness	−.001	−.049
3. Choice of suggested descriptions of mood and temperament	−.035	−.039
4. Subject's own free description of his general happiness in life	−.037	+.000
5. Adjectives applicable to subject's prevalent attitudes	−.027	−.087
6. Self-rating on satisfaction with health	−.107	−.027
7. Self-rating on satisfaction with vocation	−.144	.005
8. Self-rating on satisfaction with love and marriage	−.110	.027
9. Self-rating on satisfaction with friends	−.135	−.038
10. Self-rating on satisfaction with hobby interests	−.092	.103
11. Self-rating on satisfaction with religion	−.039	−.169
12. Self-rating on happiness in early childhood: age 3–7	−.171	−.126
13. Self-rating on happiness in later childhood: age 8–12	−.241	−.178
14. Self-rating on happiness in high school period	−.155	−.095
15. Self-rating on happiness in later adolescence	.125	−.122
16. Estimate of happiness score likely to be assigned subject by his friends	−.111	.017
17. Self-estimate of popularity with same sex in high school	−.081	−.037
18. Self-estimate of popularity with opposite sex in high school	−.010	.048

If the Otis scores be interpreted, roughly, in IQ's on the basis of the adult (eighteen year) level as given in the Otis manual they would suggest that the middle fifty per cent of this group probably fall between IQ's of 107 and 122. This, while high, is not entirely out of line with large groups of persons in other professions. While the facts for educators may not correspond to facts in other fields, still it would seem, a priori, that success in academic tests would be more likely to be related

to success and happiness among teachers than to such satisfactions in any other occupational groups one might study. If, even among teachers, it is clear that no relationship is found between test score and happiness, there is reason to be sharply critical of any assumption of general correlation. Perhaps the task of many a principal and counsellor would be made easier if it were known that the dull child and his parents might expect for that child, even in our present order, quite as much satisfaction as comes to the average, the high or the genius group.

The main portion of the study may now be considered. This consisted in the study of the relationship between total happiness and a series of factors in the experience of individuals which might be thought to have some relationship to happiness or unhappiness. An extensive questionnaire was answered by each of the 388 students who participated in the study. The questionnaire was made as objective as possible both in the matters investigated and in the control of responses. At the same time a number of subjective estimates were included, not as final determinants, but as suggestive and exploratory. This section of the study may, indeed, be regarded as preliminary effort. It attempts to break up the question and to indicate the relationships which have some possibility of being significant. Perhaps it may clear the ground of relatively unprofitable assumptions.

Table X offers a lengthy summary of some of the factors which do not appear to have any relationship to adult happiness in this group. In this table, and in several others that precede and follow it, the comparison is made between three groups of fifty each.[1] The unhappy group, the lowest fifty among the 388, range in score from 112 to 210. The group nearest the median range from 263 to 281. The fifty highest in total happiness score range from 315 to 380. The nearest extremes of the median group were thus separated by slightly more than 1PE from the happiest of the unhappy group and from the most unhappy of the happy group. The extreme groups do not come within 2PE's of any overlapping. The central tendencies of the groups are surely different. This method of comparing extreme groups was used instead of correlations, partly because it did not necessitate the assumptions of the Pearson r, but quite as much because it showed something about the direction of divergence. In some cases the factor distinguishes the happy group from both of the others. In some cases

[1] Occasionally, due to failure to answer, numbers were reduced to 49, 48 or, rarely, to 47.

the factor characterizes the unhappy group but is alike in the average and happy groups. In Table X, as has been stated, no differences between groups are significant beyond the probability that another chance selection might reverse them. Some idea of the range of items covered in the original questionnaire may be gathered from the large number of fruitless hypotheses here reported. In the original questionnaire these items were not grouped as they are here.

TABLE X.—FACTORS APPARENTLY UNRELATED TO HAPPINESS, UNHAPPINESS, OR THE MEDIAN POSITION (ALL NUMBERS ARE PER CENT EXCEPT THOSE STARRED. THESE ARE WHOLE NUMBERS)

	Unhappy Group	Median Group	Happy Group
I. *Personal Background.*			
1. Age (median)	28*	30*	30*
2. Nationality (per cent with parents not white American)	36	20	32
3. Size of town in which born:			
Under 1,000	22	22	26
1,001–10,000	24	20	24
10,001–50,000	18	24	18
Over 50,000	30	32	32
4. Size of town in which subject lived during childhood:			
Under 1,000	14	14	24
1,001 to 10,000	34	31	26
10,001 to 50,000	16	22	20
Over 50,000	36	33	30
5. Average number of physical handicaps (*e.g.*, lame, poor eyes or ears, ugly, underweight, over-weight, poor heart, digestion, lungs, etc.)	1.5*	1.4*	1.4*
II. *Family.*			
Number of children (mean)	4.2*	4.0*	5.8*
Only child	6	6	14
Middle child	32	34	32
Youngest child	26	27	34
Family somewhat more wealthy than average of community:			
During childhood	40	26	30
During adolescence	30	22	37
Family somewhat better educated than average of community:			
During childhood	54	42	46
During adolescence	44	40	40
Family demanded somewhat more obedience than average of community			
During childhood	62	56	64

TABLE X.—*Continued*

	UNHAPPY GROUP	MEDIAN GROUP	HAPPY GROUP
During adolescence	52	54	50
Family somewhat higher in social position than the average of the community:			
During childhood	54	50	50
During adolescence	49	41	42
One or both parents over forty at time of subject's birth	17	12	27
Mother had a "career" outside the home	12	10	12
Parents divorced or separated	6	6	8
Brother or sister dead	35	46	40
III. *Intellectual Factors.*			
Otis S-A Test of Mental Ability-higher Examination, average score per cent of pupils average subject believes he excelled in school marks:	47.2*	49.8*	45.5*
Elementary	82	77	77
High school	75	76	75
Normal school or college	68	75	72
Able to read own language rather rapidly	71	76	77
Able to solve algebra problems rather well or excellently	39	32	48
Read French passing well or better	41	29	39
Read Italian a little, even if very poorly	10	6	15
Read Spanish a little, even if very poorly	33	27	41
Able to explain Aristotle passing well or better	27	28	31
IV. *Extra Curricular Activities.*			
Regular editorial position on school paper or yearbook:			
High school	24	22	20
Normal school or college	16	24	16
Singing in school quartette or glee club:			
High school	34	30	40
Playing in school band or orchestra:			
High school	6	16	12
Normal school or college	10	10	12
Playing on class or practice squads for football, basketball, baseball, track, swimming, etc.:			
High school	28	26	28
Normal school or college	40	24	32
Representing school in oratory or debate:			
High school	18	18	20
Normal school or college	10	8	24
V. *Miscellaneous Abilities.*			
Dance rather well or excellently	53	43	48
Swim rather well or excellently	27	27	31
Dive rather well or excellently	12	12	13

TABLE X.—*Continued*

	Unhappy Group	Median Group	Happy Group
Play bridge rather well or excellently	37	40	43
Play poker rather well or excellently	10	7	17
Sing rather well or excellently	16	14	18
Understand radio rather well or excellently	14	12	18
Write a short story passing well or better	41	40	43
Draw a picture of someone passing well or better	20	21	14
Paint in oil a little, even if very poorly	14	17	20
Shoot at a target, average skill or better	46	44	42
Pitch a baseball, average skill or better	42	52	47
Turn handsprings, even if very poorly	23	29	33
Typewrite passing well or better	35	51	47
VI. *Membership.*			
Masons, Eastern Star, K. of C. or similar lodge	10	18	16
Political Club	2	4	6
Organization dealing with international affairs	6	14	12
Fellowship of Reconciliation, the Inquiry, or other agency for integrating social conflicts	4	8	8
American Civil Liberties Union or other free speech agency	0	2	0
American Legion, D. A. R., S. A. R., or other patriotic organization	12	8	4
Woman's Party or other feministic organization	2	4	6
League for Industrial Democracy, or other agency for reform of economic order	2	2	6
W. C. T. U., Society for the Suppression of Vice, or other agency of moral reform	0	2	2
Country Club or some similar social institution	4	12	18
VII. *Personal Attitudes.*			
Regarded in most groups as "radical" in religious ideas	33	26	40
Regarded in most groups as "radical" in educational ideas	23	18	26
Poetry one of the three things most easily spared; least essential to happiness	14	16	17
Paintings one of the three things most easily spared; least essential to happiness	24	18	14
Playing cards one of the three things most easily spared, least essential to happiness	51	55	55
Automobiles one of the three things most easily spared, least essential to happiness	4	4	12
Chance to talk with friends one of the three things most essential to happiness; one of greatest deprivations if denied	27	28	22
Studying, getting new ideas, one of the three things most essential to happiness; one of greatest deprivations if denied	18	18	15

TABLE X.—*Continued*

	Unhappy Group	Median Group	Happy Group
Good music one of the three things most essential to happiness; one of greatest deprivations if denied	18	10	18
Novels and stories, one of the three most essential things to happiness; one of the greatest deprivations if denied	12	10	6
Preferring to see a play rather than act in it	85	86	83
Regarded by other children in community as unusually good, pious, studious, inclined to be a teacher's favorite	37	30	35
Have had strong emotional attachments to persons of same sex	21	24	14
Not easily shocked	94	98	96
Believe sex education was wisely given	37	39	47
Estimate that they exceeded half their high school classmates in popularity with fellow-pupils of the opposite sex:			
75 men	42	42	55
126 women	29	36	38

A number of surprising findings merit special comment. Age, ranging in these groups from under twenty to over sixty seemed to be unrelated to happiness. Apparently the best of life is not necessarily found either first or last. Bringing children up in small towns rather than in large cities seems to be no guarantee of greater adult happiness. That possession of physical defects should not have contributed to unhappiness still seems surprising. While family size was in general large, even for thirty years ago, it is still interesting to note that there is no clear tendency for the larger families to be more frequent in the happy group. Neither are only children unhappy adults. While the cases are few, the evidence tends slightly in the opposite direction. The writer was surprised to discover no clear tendency for those who as children were brought up with more demand for obedience to be less happy as adults. A previous study had connected shyness with obedience demanded in childhood, and unhappiness seemed a fair inference. The old bugaboo about having children when parents are past forty does not seem to be justified, when results are measured in terms of child happiness. The mother's career is likewise not a factor to be deplored, from the standpoint of the children's later welfare. The cases involving divorce or separation are few, but suggest that this is not a significant factor. Harmony between parents did appear significant, as will be shown later. The negative results on

school subjects was to have been expected by everyone, presumably, except the defenders of the present curriculum. There are those who pretend (whether at heart they really believe it I never could quite decide) that life will be richer and happier for those who know better how to read efficiently, who know languages, philosophy and mathematics. It is refreshing to note how little support, even in the academic group itself there is for such assumptions. If school marks count anywhere, they should indeed be useful for teachers. To former studies showing little correlation between teaching efficiency and school marks, may now be added the implications of this finding. There seems to be no correlation between life satisfaction among educators, and their school mark achievements. It was more surprising to find that the extra-curricular activities likewise contributed so little. Perhaps the missing element here is the quality of performance in these extra-curricular fields. At present, the facts point toward no clear relationship between later life happiness and having taken part in editorial, musical, or athletic activities in high school or college. The quality of performance enters into the judgments of the next section on special abilities. Subjects could rate their ability from "Can't do at all," through grades, "Do but very poorly," "Do passing well," "Do rather well" to "Do excellently." Some abilities as will later be noted did seem significant, but athletic, social, and artistic abilities seem here to be uncorrelated with happiness. It is perhaps, not surprising that the organizational memberships should matter so little. There seemed to be few of them, anyhow.

In the personal attitudes section it is interesting to note that the religious radicals were not unusually common among those unhappy in life. If anything they were found more often in the happy group. There is a slight tendency for this distribution to support Allport in saying that extremists, at both ends of the scale, are more alike than are the moderates like either extreme. The moderately happy were not strong on religious radicalism. The extremes were somewhat more so, although the differences are too small to count upon. The writer was again surprised to find that having been regarded by other children in the community as unusually pious and a teacher's pet did not in any sense doom the child to later unhappiness, at least if that child enters the teaching profession. Crushes appeared likewise with equal frequency in the happy and unhappy groups. It was astonishing and almost unbelievable that those who could report their sex education wisely given were not significantly more often

found in the happy group than in the unhappy group. There was a difference in this direction, but it was too small to regard as significant. Popularity with the opposite sex was another factor, which while tending in the direction of happiness was so small as to make it necessary to regard it as insignificant. This, it will be remembered, was the person's own estimate of his popularity with the opposite sex. More objective measures of popularity later show that this factor is more important. Nevertheless, it is still not clear why individuals believing themselves popular should be found almost as often in the unhappy group as in the happy group.

In Table XI we come upon the main positive findings of the study. Here are presented the factors showing a significant difference between these three groups. The table shows not only the per cent found in each group, but the difference between happy and unhappy with the significance of that difference. Only the first fourteen factors may be said to differentiate between the groups to the satisfaction of the statistician. The remaining factors showing a difference larger than its SD are presented, however, in the table. There are two reasons for doing this. One is that a difference only twice its SD still has ninety-six chances out of a hundred of recurring in the same direction with other samplings. These are respectable odds. A difference only 1.5 times its SD may still be expected to lie in the same direction in eighty-seven per cent of future samplings. The second reason for presenting the less certain differences is that they may have a cumulative effect. One item alone may be a bit uncertain. A constellation of items, pointing in the same general direction, has much more to warrant attention. The reader may, therefore, wish to glance over the less certain differences to note several which reinforce each other.

No one would probably have hit in advance upon the one of the several hundred items which proved most significant in differentiating the happy and unhappy groups. Apparently normal and happy groups divide about 60-40, in desiring the company of some highly desirable companion of the opposite sex rather than that of some person of the same sex. In the unhappy group, however, ninety-five per cent chose the opposite sex. It would look as though disappointment in love relationships could be said to be a major factor in unhappiness. Other factors, such as the difference in per cent married, (eight per cent instead of thirty-six per cent), the demand for parties and dances among the unhappy, the broken engagements (sixty-eight

TABLE XI

Factor as reported	Per cent reported by			Difference between happy and unhappy	Difference / SD difference
	50 Unhappy	50 Median	50 Happy		
Differences between Happy and Unhappy Groups[1]					
1. Prefer to spend a few hours with most delightful known companion (not in the family) of same sex, rather than opposite sex	5	40	42	37	4.60
2. Able to give a talk to high school pupils about sex, passing well or better. All:	27	50	69	42	4.55
75 men	31	57	96	65	5.3
126 women	26	36	48	22	2.2
3. Able to administer a large group of workers rather well or excellently	21	27	58	37	4.05
4. Not particularly sensitive to opinion of others	16	37	50	34	3.84
5. Not bothered by strong fears	65	92	94	29	3.81
6. Health during adolescence, superior	26	32	60	34	3.66
7. Able to repair a fuse rather well, or excellently	8	30	35	27	3.44
8. Not shy, timid or unduly self-conscious	49	72	80	31	3.40
9. Tend to prefer a job with great responsibility, rather than one which leaves the individual rather free, others carrying the responsibility	51	76	82	31	3.35
10. Able to lead an effective group discussion, passing well or better	51	75	81	30	3.28
11. Able to make an acceptable, or better impression at a tea	73	90	96	23	3.23
12. Married	8	13	36	28	3.04
13. Not preferring to live alone at age sixty	75	87	96	21	3.02
14. Believe love and harmony with mother at or above average for community:					
126 women (during childhood)	66	34	91	25	3.0
126 women (during adolescence)	64	80	82	18	1.96
Differences Significant from Two to Three Times Their Standard Deviation					
Able to tell a joke to a crowd, acceptably or better	51	75	78	27	2.92
Not apt to get tired very quickly	56	84	82	26	2.90
Dances and parties among the most easily spared activities, among the three least essential to happiness	8	29	28	20	2.76

[1] Factors are arranged in order of the statistical reliability of the obtained difference. The more eccentric group is underlined to indicate whether the factor is more clearly related to happiness or to unhappiness.

TABLE XI.—*Continued*

Factor as reported	Per cent reported by			Difference between happy and unhappy	Difference / SD difference
	50 Unhappy	50 Median	50 Happy		
Love and harmony between parents at least average for community during subject's childhood	68	84	90	22	2.70
Superior health during childhood	24	36	48	24	2.59
Subject's work among the three activities most essential to happiness, greatest deprivation if denied	51	75	75	24	2.54
Able to act in a play passing well or better	49	51	73	24	2.51
Held two or more elective offices in normal school or college	36	42	60	24	2.48
Male (sixty-four per cent of men exceeded median woman)	14	28	34	20	2.39
Having been engaged, are now engaged, married, widowed, or divorced (*i.e.*, no broken and unreplaced engagement)	32 (of 28)	60 (of 22)	61 (of 23)	29	2.37
Able to play tennis passing well or better	29	47	53	24	2.36
Able to make out a rather good or excellent questionnaire	15	25	34	19	2.24
Member of teacher's federation, labor union, etc	10	18	27	17	2.22
Able to do rather well or excellently at rapid, accurate addition	27	33	48	21	2.20
Preferring to think things through rather than to meet situations impulsively, doing what seems appropriate on the spur of the moment	82	92	96	14	2.19
Preferring at age sixty, to be living with husband or wife	47	38	68	21	2.14
Unable to write a rather good or excellent poem	75	94	91	16	2.12
Some ability to do modeling and sculpture, even if very poorly	10	21	26	16	2.11
Clubs not easily spared, not among the three least essential things for personal happiness	27	37	47	20	2.10
Alcoholic liquor most easily spared, among the three things least essential for happiness	77	92	92	15	2.10
Unable to play the piano, or able to play only very poorly	59	71	78	19	2.06
Spent years from fourteen to twenty-one in community of less than 1000 population	4	8	11	12	2.04
Not preferring, at age sixty, to live with a sister	87	94	98	11	2.02

TABLE XI.—*Continued*

Factor as reported	Per cent reported by			Difference between happy and unhappy	Difference / SD difference
	50 Unhappy	50 Median	50 Happy		
Differences Significant to From One to Two Times Their Standard Deviation. No Differences Included Unless the Per Cent in One Group Exceeded the Other by More than Ten					
Father dead	29	36	48	19	1.97
Some ability to identify a myrtle warbler	18	29	35	17	1.94
Love and harmony between parents during subject's adolescence at least average for community	71	78	88	17	1.93
Ability to read Latin passing well or better					
Churches and religious activities not so easily spared, not among the three least essential for personal happiness	55	78	73	18	1.89
No serious disappointment in efforts to get grades or other scholastic honors	85	90	96	11	1.89
Ability to distinguish the various writers of the Pentateuch, passing well or better	10	16	24	14	1.88
Preferring at age sixty, to live with friends of the same sex	6	13	19	13	1.87
Able to classify a new flower passing well or better	22	30	39	17	1.86
Took part in play in college or normal school	34	38	52	18	1.85
Held two or more elective offices in high school	38	38	56	18	1.83
Now member of hobby club, musical, literary, or other special-interest group	24	29	41	17	1.83
Much interested in religion	37	48	55	18	1.81
Not an oldest child	64	67	80	16	1.81
Worked to help support self while in school	31	53	48	17	1.75
Able to read German passing well or better	14	24	28	14	1.74
Received at least the average amount of attention from average sex during adolescence	67	80	82	15	1.73
Now member of some social fraternity or social club	43	45	61	17	1.72
Able to understand technical scientific articles at least passing well	63	69	78	15	1.67
An only or youngest child	32	33	48	16	1.66
Family, during subject's childhood, more interested in religious activities than was the average family in the community	46	48	62	16	1.61
Estimate having exceeded half classmates in popularity among high school fellow-pupils of same sex:					

TABLE XI.—*Continued*

Factor as reported	Per cent reported by			Difference between happy and unhappy	Difference / SD difference
	50 Unhappy	50 Median	50 Happy		
75 men	40	50	67	27	1.65
126 women	53	66	70	17	1.61
Mother dead	24	26	39	15	1.61
Prefer to think of life as earnest, worthwhile, to be lived, with serious purposes rather than as amusing, or to be taken lightly	78	83	90	12	1.61
Sung in college glee club, quartette or choral society	22	32	36	14	1.54
Getting out of doors, enjoying nature, one of the three things most essential to happiness, one of greatest deprivations if denied	39	37	55	16	1.47
Being able to skate passing well or better	45	39	59	14	1.40
Family during subject's adolescence, somewhat more interested in religious activities than was average of community	44	47	58	14	1.36
Prefer life adventurous, changing or exciting, rather than quiet, serene, free from worry and settled	67	68	79	12	1.31
Participated in class play in high school	44	52	56	12	1.21
Religion not much like that of parents	37	44	48	11	1.11

per cent to thirty-nine per cent) and the smaller per cent receiving at least the average amount of attention from the opposite sex during adolescence (sixty-seven per cent to eighty-two per cent) would seem to confirm the importance of this item.

Second in significance stands an item which combines the factor of sex maladjustment with that of self-consciousness and a sense of inadequacy. The actual differences between groups are larger on this than on any other item studied. Subjects were asked how well they could do at giving a talk on sex education to high school pupils. Sixty-nine per cent of the happy group answered, "Passing Well," "Rather Well," or "Excellently." Less than half as many of the unhappy group (twenty-seven per cent) were so confident. Aware that men were found more often in the happy than in the unhappy group it seemed wise to note possible sex differences in this answer. With the sexes separated, however, the item is still very diagnostic. It is particularly so with the men, ninety-six per cent of the happy group feeling adequate for such a responsibility, while less than a third of the unhappy men felt they could do well in such a situation.

The factor of adequacy stands out in the third item, confidence in one's ability to administer a large group of workers. This may fairly be said to represent a pattern which distinguishes the happy and unhappy groups. Not only are the happy group better able to give talks to high school pupil, to administer workers, but they rate themselves as more adequate, also, in leading group discussions, making acceptable impression at a tea, telling jokes to a crowd, or acting in a play. As a matter of fact they have held more offices, acted in more plays, belonged to more fraternities, and they tend, in greater proportion (eight-two per cent to fifty-one per cent) to prefer a job with great responsibility rather than one which leaves the individual free while others carry the load. One is reminded of Chassell's finding[1] that enjoyment of social intercourse and freedom from hyper-self-consciousness are associated with general happiness and satisfaction in life to the extent indicated by coefficients of association (Q) of .43, .57, .72, or .74.

Before this interpretation is accepted, however, one other hypothesis should be examined. Perhaps the happy group are generally reckless and extreme in their oversatisfaction with themselves. Perhaps they overestimate all their abilities, as well as their comparative satisfaction in life. Perhaps the unhappy group are reluctant to call themselves happy or to admit the abilities they really have. One answer to this suggestion is found, of course, in the factors which failed to differentiate. It will be remembered that in Table X there were many more items than are here reported as showing a difference, upon which no difference in self-estimates of ability appeared. Some of these concerned school subjects, others athletic or musical prowess. Table XII, below, sums up the situation for all the forty-three abilities studied. There was obviously no clear tendency for the happy group to claim high ability in general, or for them to avoid the extreme of inability.

The only conclusion justified would seem to be that there is a real difference in the tendency of the happy group to think of themselves as adequate in situations requiring making a satisfactory impression upon other people.

There are other less important factors which are interesting enough to warrant special comment. The unhappy group is particularly characterized by sensitiveness and fears, timidity, and self-conscious-

[1] Chassell, J. O.: "The Experience Variables." Published by the author, 1088 Monroe Avenue, Rochester, N. Y., 1927.

ness. Love and harmony with the mother, the father, and between the parents, was greater in the homes from which the happy group come. Superior health during childhood (forty-eight per cent to twenty-four per cent) and adolescence (sixty-six per cent to twenty-six per cent) was markedly characteristic of the happy individuals. It would appear safer to judge an individual's present happiness by his health during earlier years than by his happiness during earlier years. The larger proportion of men and married persons in the happy group suggests that this may be still a man's world, and that whatever

TABLE XII.—PER CENT OF EXTREME RESPONSES ON FORTY-THREE JUDGMENTS OF ABILITY (BASED ON A FIVE POINT SCALE)

	Can't do at all (per cent)	Do excellently (per cent)	Total extreme (per cent)
Unhappy	36	5	41
Median	32	4	36
Happy	30	6	36

is wrong with marriage it still is a good competitor with single blessedness. Interesting, rather than significant is the somewhat larger proportion of unhappy individuals who believe themselves able to write a rather good poem (twenty-five per cent to nine per cent), to play the piano at least passing well (forty-one per cent to twenty-two per cent), to draw an acceptable picture of someone else (twenty per cent to fourteen per cent). Not only the arts, but their traditional ally, alcohol, seems to furnish balm for the unhappy. Alcohol was placed among the non-essentials by ninety-two per cent of average or happy persons but by only seventy-seven per cent of the unhappy group. Religion escaped classification with the anodynes, due perhaps to the presence in the group of a number of specialists in religion and religious education, working from the liberal point of view of Union Theological Seminary. Churches and religious activities were, at any rate, non-essentials for forty-five per cent of the unhappy group but for only twenty-seven per cent of the happy group. Only thirty-seven per cent of the unhappy students but fifty-five per cent of the happy group report themselves much interested in religion. Forty-six per cent of the homes of unhappy people and sixty-two per cent of the homes of those who regard themselves as happy were more

interested in religion than was the average member of the community. The somewhat liberalized religious interest is indicated perhaps, by the fact that only fifty-two per cent of the happy group and sixty-three per cent of the unhappy group report religious ideas much resembling those of their parents. A further indication was that knowledge of the several writers of the Pentateuch characterized the happy group (twenty-four per cent) rather than the unhappy group (ten per cent).

Order in the family showed a slight difference, only or youngest children predominating in the happy group (forty-eight per cent to thirty-two per cent) and oldest children in the unhappy group (thirty-six per cent to twenty per cent). This agrees with several recent studies showing that the lot of the only child has been regarded with unnecessary sympathy, and that lack of aggressiveness is particularly the problem of the oldest youngster.

A suggestion that love of nature plays its part in happiness is found in the fact that more of the happy group passed their adolescence in rural communities (sixteen per cent to four per cent), more can identify a myrtle warbler (thirty-five per cent to eighteen per cent), or classify a new flower (thirty-nine per cent to thirty-two per cent). In the happy groups, also, there was a more marked tendency to put "getting out of doors, and enjoying nature" as one of the essentials of happiness (fifty-five per cent to thirty-nine per cent).

If there be any trend toward the flippant life among youth today, it might be well to ponder the predominance of serious interests and purposes in the happy group. Dances and parties were more important for the unhappy. The happy more markedly enjoyed their work, preferred jobs with responsibility, preferred thinking things through to acting on impulse, worked to support themselves in school, were able to read Latin, German or technical articles, and preferred to think of life as earnest, worth-while, to be lived with serious purpose, rather than as amusing or to be taken lightly. On the other hand, the preference of all three groups, but particularly of the happy group, leaned toward a life that was adventurous, changing, and exciting rather than one of quiet, serenity, and security.

The factor of popularity is certainly related to happiness. The happy individuals, as has been suggested, were less sensitive and shy, more apt to hold elective offices, more apt to avoid broken engagements, enjoy clubs more, more apt to have received at least the average amount of attention from the opposite sex during adolescence (eighty-

TABLE XIII.—CORRELATIONS WITH SELF-ESTIMATES OF POPULARITY IN HIGH SCHOOL

Correlated with	Men, popularity with		Women, popularity with	
	Same sex	Opposite sex	Same sex	Opposite sex
1. Total happiness score	.341	.332	.200	.084
2. Subject's self-estimate of general happiness	.437	.211	.096	.073
3. Choice of suggested descriptions of prevailing mood	.028	.125	.116	.030
4. Original free description of general happiness in life	.220	.392	.098	−.093
5. Check of adjectives	.237	.150	.140	−.017
6. Self-rating on satisfaction with health	.371	.100	.116	.301
7. Self-rating on satisfaction with vocation	.180	.089	.230	.031
8. Self-rating on satisfaction with love and marriage	.366	.328	.204	.167
9. Self-rating on satisfaction with friends	.292	.245	.250	.030
10. Self-rating on satisfaction with hobby interests	−.077	−.103	.186	.079
11. Self-rating on satisfaction with religion	.011	.009	.013	−.044
12. Self-rating on happiness in early childhood (3–7)	.337	.102	.167	.073
13. Self-rating on happiness in later childhood (8–12)	.173	.133	.180	.077
14. Self-rating on happiness in high school period	.295	.113	.181	.076
15. Self-rating on happiness in later adolescence (18–25)	.258	−.069	.191	.063
16. Estimate of happiness score likely to be assigned subject by his friends	.244	.304	.091	.009
17. Otis Intelligence Score	−.081	−.010	−.037	.048
Correlation of self-estimates of popularity, with same sex, with self-estimates of popularity with opposite sex	.641		.357	

two per cent to sixty-seven per cent) and more apt to estimate their popularity as exceeding their average classmates. Estimates of popularity with the same sex, however, proved better at differentiating than did estimates of popularity with the opposite sex. Table XIII studies the problem using the correlation technique. It appears that the correlations are not high, those of the estimates by women of their popularity with the opposite sex being apparently unrelated to any other factor. This would be worth some additional study of a more objective sort. Is the discrepancy in the modesty of the popular girls, the pride of the unpopular girls, or is the relationship really negligible? The significant correlations with other indices of popularity and sex satisfaction point to the conclusion that women's estimates of their popularity with the opposite sex are not to be trusted.

TABLE XIV.—PER CENT OF EXTREME RESPONSES ON SIXTEEN JUDGMENTS CONCERNING SUBJECT'S FAMILY STATUS DURING CHILDHOOD AND ADOLESCENCE (BASED ON A FIVE POINT SCALE)

	Per cent extreme deprivation	Per cent extreme advantage	Per cent total extreme
Unhappy	10	18	28
Median	4	17	21
Happy	4	22	26

There remains the possibility of factors distinguishing the median group from both extremes. Table XII has already shown that the group rating themselves nearest the middle of the group in happiness, were not equally moderate at all points. The per cent of extreme judgments upon ability, either positive or negative, was the same for the median group as for the very happy group. Table XIV gives the summary of sixteen judgments in which the subject indicated, on a five point scale, how his family compared with other families of the community in wealth, education, social position, love and harmony, interest in religion, and so on. It appears that the median group were not markedly more temperate in their estimates on these matters than were the extreme groups. They used the extreme categories for twenty-one per cent of their answers as compared with twenty-six per cent for the happy group and twenty-eight per cent for the unhappy group. This difference might easily be due to chance or to a real difference in environment since slightly more social deprivation

appeared with the unhappy group and slightly more advantage with the happy students. In other words, there is no warrant for assuming that the median group is characterized by moderation and caution in all answers. It was specifically the degree of happiness which

TABLE XV.—FACTORS DISTINGUISHING THE MEDIAN GROUP FROM BOTH HAPPY AND UNHAPPY EXTREMES

Factor as reported	Per cent reported by			Difference between median and average of extremes	Difference
	Unhappy	Median	Happy		SD difference
Connected with a church or synagogue	65	92	73	23	3.79
Never have been homesick for as much as a week	55	80	64	20	3.54
Never represented college on an athletic team	70	88	70	18	2.77
Never represented high school on an athletic team	66	84	66	18	2.57
Connected with missionary and other religious organizations (other than churches)	14	43	31	21	2.55
Not regarded in most groups as radical in social and economic ideas	71	80	60	15	2.42
Theatre not any of the three things most essential to happiness; not one of greatest deprivations if denied	79	94	86	11½	2.29
Never have been engaged to be married	43	55	39	17	1.97
Making things with hands not among the three activities most easily spared; not among those least essential to happiness	78	88	76	11	1.75
Member of charity organization, relief society, society for aiding blind, deaf, orphaned, etc	10	24	14	12	1.73
Able to repeat the Twenty-third Psalm, passing well or better	65	76	62	12	1.67
Dance not at all or very poorly	11	27	18	12	1.67
Member of professional club, fraternity or society	49	67	59	13	1.55
Member of Red Cross	61	73	63	11	1.37

placed them between the unhappy group on the one hand and the happy group on the other.

Table XV reports the factors showing the moderately happy group exceeding both the unhappy and the happy. Their most marked characteristic is participation in churches, synagogues, missionary or other religious societies, charity organizations, professional societies, and the Red Cross. There is a slight tendency for them to be unusual in never having gone in for athletics, radical ideas, or matrimonial plans.

TABLE XVI.—SCORE OF SIXTEEN ACTIVITIES SHOWING EXTENT TO WHICH THEY WERE BELIEVED MOST ESSENTIAL FOR HAPPINESS

Activity	Total group		Unhappy		Median		Happy	
	Rank	Score	Rank	Score	Rank	Score	Rank	Score
Chance to talk with friends	1	252	1	82	1	97	2	73
Their work	2	228	2	57	2	81	1	90
Studying, getting new ideas	3	147	3	56	3	47	3	44
Getting out of doors, with trees, flowers, birds, lakes, etc	4	88	4	27	4	23	4	38
Good music	5	36	5	18	5	7	6	11
Theatre	6	30	6	15	7	2	5	13
Novels and stories	7	15	7	9	6	3	7	3
Automobiles	8	−8	9	−3	8	1	8	−6
Poetry	9	−36	10	−9	10	−10	11½	−17
Paintings	10½	−37	11	−13	12	−17	9	−7
Dances and parties	10½	−37	8	1	13	−20	13	−18
Chance to make things work with their hands	12	−40	12	−19	9	−4	11½	−17
Churches and church activities	13	−64	14	−38	11	−15	10	−11
Clubs	14	−80	13	−33	14	−26	14	−21
Playing cards	15	−149	15	−47	15	−48	15	−54
Alcoholic liquor	16	−354	16	−103	16	−124	16	−127

One of the sections of the questionnaire asked students to choose from a list of sixteen activities the three which were most essential to happiness and the three which could be most easily spared so far as they were concerned. The results appear more useful in showing the

agreement among all of the students than in differentiating particular groups. Table XVI shows the rank of the activities for the total group, the unhappy, median, and happy groups. Actual scores as well as ranks are given. A negative score signifies that the activity was chosen more often as one that could readily be spared than as one which was essential to happiness.

Among the outstanding differences, the higher rank given to dances and parties by the unhappy group, the higher rank given to paintings by the happy group and to hand work by the median group are the only significant comparisons. In general, however, it is instructive to observe the high place given to friendship, work, intellectual stimulus, and nature. It may be a bit surprising to find nature ranked above all the arts. Churches, on the other hand, are negative for all three groups.

In conclusion, some of the more challenging findings may be brought together in the form of hypotheses. These will be stated as tersely as possible, and hence appear more dogmatic than the selected group upon which they are based would warrant.

1. Intelligence has no relation to happiness.
2. Failure in love is a major cause of unhappiness.
3. Enjoyment of, and success in, work is a major factor in happiness.
4. Good health in childhood is the foundation for happiness.
5. Popularity matters.
6. School marks do not matter.
7. Success in dealing with people is fundamental to happiness.
8. Music and poetry tend to be refuges for the unhappy.
9. Religion, of the modern type, is not merely an escape for the unhappy.
10. Graduate students of education are, on the whole, fairly well satisfied with life.
11. General level of happiness among such adults can be measured with adequate reliability by a single check on one graphic scale.
12. The unhappy believe they give an impression of greater happiness than they feel.
13. Hobbies are not so important for life satisfaction as is sometimes supposed.
14. Youth is not the golden era of happiness; neither is age.
15. Only children are as likely, perhaps more likely, to be happy as are children in large families.
16. The comparative wealth of parents does not affect the happiness of children.
17. Educated parents do not raise happier children.
18. Quarrelsome parents, divorced, seem to hurt a child's happiness less than the same parents remaining together.
19. Mother careers do not make for unhappiness in children.
20. More or less knowledge of academic subject-matter does not influence happiness.

21. Participation in athletics is not significant for or against later happiness.
22. Parents over forty at the time of the child's birth are as likely as any to raise happy children.
23. Ability at dancing, cards, athletics, writing, music, or painting is unrelated to happiness.
24. "Radicals" are not the fundamentally unhappy.
25. A history of crushes does not necessarily indicate unhappiness.
26. A sex education, now regarded by educated adults as "wise," did not improve chances for happiness.
27. Those who look forward to living at age sixty, alone or with a sister, are predominantly unhappy.
28. The married are happier than the unmarried.
29. Men believe themselves happier than women believe themselves to be.
30. Success in dramatics is more indicative of happiness than is success in any other form of extra-curricular activity.
31. Blessed are those who are elected to many offices.
32. Love of nature goes with greater happiness.
33. The essentials of happiness for most people are among the stable elements of life (friends, work, nature) not among the stimulants (alcohol, clubs, churches, dancing, cards, automobiles, or arts).
34. Most people crave adventure rather than serenity.
35. Creative work with the hands is not fundamental to happiness.
36. Fears, sensitiveness, shyness, are rightly regarded as major factors in unhappiness.
37. Happiness is associated with serious, deliberate, responsible, earnest, hard-working living, rather than with impulsive, light, amusing, dilettante-ism.
38. The proportion of error in these statements is uncertain, but considerable.

A TRUCE TO "BARKING IN"

C. SPEARMAN

University of London

Being deeply interested in the subject—and moreover noticing mention of my name—I turned with expectation to the recent work of Mr. H. D. Carter conducted under the direction of Professor D. G. Paterson at the University of Minnesota.[1]

The commencement of the work proved disappointing. Two-thirds of it were found devoted to the "inter-columnar correlation" as criterion for the presence of the familiar "two factors," *g* and *s*. This correlation is well known to have long been quite obsolete; indeed, even originally it had only been put forward as a stop-gap pending discovery of the correct criterion. Moreover, in order to be usable at all, even as a stop-gap, this correlation was obliged to fulfil a certain condition.[2] Now, as regards one of his two tables of data, this condition is not mentioned by Carter at all; as regards the other table, it is mentioned and said to be fulfilled for every column. In truth, unfortunately, it is not fulfilled in either table; not nearly!

After this unhappy start, however, he makes a good recovery. Employing at last the correct criterion (that of "tetrad differences"), he brings evidence that, when some tests of "general intelligence" or *g* are put into company with tests of "mechanical ability," the criterion fails; the mechanical abilities manifest what is called a group factor.

But why does he then go on to turn this finding of his into a polemic against the theory of two factors as held by my colleagues and myself? If he will look again at my book, "The Abilities of Man," which he chiefly quotes, he will easily find in it (by using the index) the original demonstration of this very fact about mechanical ability—which he is now the fourth, I think, to re-discover! He will also find much relevant to it, including an exposition of its theoretical significance.

As another type of critic, I feel that a word should be said about the recent work of Professor W. F. Dearborn.[3] He makes many quotations from the same book of mine. But one is surprised to note

[1] *Journal of Genetic Psychology*, 1928.

[2] This to the effect that "the mean square probable error of the coefficients in any column should not be greater than a quarter of the mean square deviation of those coefficients." *British Journal of Psychology*, 1912, p. 56.

[3] "Intelligence Tests," 1928.

that, although he speaks in a markedly antagonistic emotional tone, he does not even attempt any antagonistic argument. He brings everything to inflame the combat, but nothing to help it; nothing that can be verified by further observations, or even clarified by further discussion; nothing, in a word, which makes for any scientific progress. Still stranger is it that all his quotations and references come from the smaller and controversial part of my book (the part entitled "The Rival Doctrines"). As for the far larger part which as its title shows ("The Fundamental Facts") is not devoted to controversial theory but sets forth a great accumulation of actual observations fundamental for his own work with all this he shows no sign of having concerned himself.

The behavior of such authors is all too suggestive of our canine friends when they chance upon a street row. With fine scorn for the actual matter in dispute, they joyously "bark in," if the expression may be pardoned me. But exhilarating and easy as this no doubt is, we must remember that it has been going on for a great many years. And just now the opposing parties seem to have arrived at the stage of understanding each other well enough to bring the chief relevant facts into serious consideration.[1] Might we not be wise to take advantage of it?

[1] See "The Psychologies of 1930."

FURTHER STUDIES OF CHILDREN'S INTEREST IN COLLECTING

PAUL A. WITTY

University of Kansas

AND

HARVEY C. LEHMAN

Ohio University

Miss Mary T. Whitley has recently reported some findings regarding children's interest in collecting which are in sharp contrast to some data reported previously by the present writers. Those interested in a detailed account of the facts leading up to the present discussion are referred to the accompanying citations.[1,2] For the present purpose it need only be said that, incidental to their study of children's play behavior, the present writers found that less than thirty per cent of 5000 subjects reported at any season of the year an active interest in making collections. At most age levels, the per cent was much smaller than this. After use of a quite different and somewhat ambiguous technique, C. F. Burk asserted many years ago that ninety per cent of the boys and ninety-one per cent of the girls studied by her were actively making collections.[3]

As a result of a recent investigation, in which she employed a technique differing markedly from that of Burk or that of Lehman and Witty, Miss Whitley secured data which differ in certain respects from those obtained by her predecessors. Since quite different techniques were utilized in all three studies, it is natural to expect that the results of the three investigations would not agree closely. Undoubtedly, the results of a given investigation are, to a large extent, a function of the procedure employed for obtaining them.

At this point the writers wish candidly to admit the justness of certain of Miss Whitley's criticisms of their conclusions regarding children's interest in collecting. Repetition of studies has justified itself in more than one instance in the history of psychology. Moreover, it will be admitted readily that the concentration of several available techniques should yield valuable data regarding many elusive problems of human behavior.

The Problem of Terminology

The writers feel that a word or two with reference to ambiguity of terminology is very pertinent at this point. The question of

terminology is not an idle one. Indeed, the problem of definition is in their judgment no less important than the problem of technique. Burk reported that approximately ninety per cent of her subjects were actively making collections at the time of her investigation. Miss Whitley shows that no less than eighty-five per cent of her subjects were interested in making collections. Here one must pause and inquire what is meant by "collecting." Greater precision in definition of terms than the subject requires is mere pedantry; but, until there is some agreement regarding the meaning of the term "collecting," the reports of all writers will remain ambiguous.

Burk presented a list of nearly three hundred items which her subjects (all children) reported that they had collected. Miss Whitley reports that, including the sixty articles in her check-list, her returns totaled three hundred seventy-eight articles (after classifying many together, such as separate articles of clothing). Miss Whitley does not present the names of all the three hundred seventy-eight articles, but examination of Burk's list reveals some novel interpretations of what it means to "make a collection."

Miss Whitley frankly states that the child has difficulty in comprehending the meaning of the word "collection." Perusal of Burk's list of nearly three hundred items suggests to the writers that the children are not alone in their misunderstanding of the term. Just what *does* it mean to "make a collection?"

Several investigators have asserted that "collecting and hoarding" is instinctive or an instinct, that collecting is an innate (unlearned) tendency to respond in a definite way to certain stimuli. The child is assumed to collect "portable objects" which attract attention just for the love of the activity itself. The activity is motivated supposedly by inner urges which are the result of inheritance.

> There is originally a blind tendency to take portable objects which attract attention, and carry them to one's habitation. There is the further response of satisfaction at contemplating and fingering them there.[4]

The above citation assumes the innateness of the collecting tendency. Norsworthy and Whitley discuss this so-called instinct and conclude that

> The strength of this tendency (collecting) in childhood and the fact that it is still present in so many adults suggest that the school would do well to use the *instinct* more.[5] (Italics ours.)

Statements similar to the last cited one are found in most approved texts in psychology. Even R. M. Ogden in his attempt to provide

the Gestalt approach in education stresses the importance of "this instinct" (collecting).[6]

To anyone familiar with the present array of theories in regard to the instinct hypothesis, *one conclusion only seems warranted.* Many of the items listed by Burk are clearly not prompted by an hereditary predisposition to collect but by environmental stimulations which are easily discernible.

Burk's list of things collected included many items which may have been merely noticed or counted. Among these were white horses, red-headed girls, bald-headed men, hat-tips, etc. According to such a concept of collecting, a man possessed by a strong compulsion to count cracks in a cement sidewalk would be a "collector" of cracks-in-the-cement-sidewalk.

What sorts of things make a collection? If a child obtains bits of information, (*e.g.*, in the field of history), is it legitimate thenceforth to refer to the child as a "collector" of historical information? If this be granted, it must follow that *every* normal individual is a collector, for certainly every individual acquires innumerable bits of information.

In Burk's list it is reported that some children were collectors of "bits of advice," "knots in ropes," "riddles," "witty sayings," "all sorts of sayings," "stories," etc.

Just what would be the observable activity of a child given to collecting "bits of advice," or "knots in ropes?" Would the latter of these two activities mean that the child has learned or was learning how to tie various kinds of knots in ropes? The writers are anxious to know whether the term "collect" has any clear connotation whatsoever. Does it mean anymore than the ambiguous term "instinct?"

The above are only a few of the items of Burk's list of objects collected which the present writers find to be extremely ambiguous as well as ridiculous. The writers have confined their criticism chiefly to Burk's list because they have not seen Miss Whitley's list of three hundred seventy-eight items. The sixty items included in Miss Whitley's check-list lead them to no lessened conviction that the meaning of the term "collection" needs clarification. Included in the Whitley check-list are such terms as "samples of school work," "letters received," "photographs," "old magazines," "theater programs," etc.

The problem of definition is exceedingly important. If "to make a collection" means merely an unwillingness to throw objects away or if it means merely to allow objects to accumulate, this fact should be stated clearly in order that the reader may know what is meant

when returns are expressed in per cents; *e.g.* it makes a vast sex difference in the returns if we include under "collecting" the retention of letters received, for girls appear to receive and to write letters much more frequently than do boys.

Burk states that forty-eight per cent of her girls were "given" the objects which comprised their collections. Such a report gives a quite different connotation from that concerning a boy who is aggressively adding to his collection by earnest effort. Until some definition of terms is agreed upon it will be impossible, of course, to make comparisons of the results of studies of the general tendency to "collect."

Another troublesome question arises concerning the definition of collecting: *How many* objects make a collection? What is the *minimum* number of objects that may comprise a collection? If a man purchase a dog or two for his children would this be legitimate grounds for reporting that one of his children is now a "collector" of dogs? Probably not, yet, according to Burk's list, some of her subjects reported that they were collecting goats, monkeys, cows, horses,* and various other animals. Just what would be the observable activity of a child engaged in collecting cows or horses? Certainly, the concept which Burk must have held could hardly be accepted even by the "instinct" psychologists.

When children report that they engage in collecting edibles, what are we to suppose they mean? Is it to be supposed that the children assembled these edibles merely for the purpose of consuming them? The writers ask not in a carping spirit but as earnest seekers of information. What does a child mean (if anything) when he reports that he is collecting "candies," "cookies," etc.? These items are included in Burk's list.

It is of interest that Burk lists five methods by which her subjects added to their collections. These five methods are: "Finding," "Given," "Buying," "Trading," "Winning." Are all of these methods of collecting a single form of behavior? Obviously not. Are they five different aspects of an instinct to collect? This hypothesis is also an extremely doubtful one.

Presumably, a boy is "collecting" when he purchases or is given stamps to add to his collection. The boy reports also that he is "collecting" when he assembles bottles to sell to the junkdealer.

---

* This report referred evidently to some activity other than the counting of white horses.

Are these two types of behavior prompted by a single instinct or by the same instinct?

Woodworth has suggested that the tendency for the child to take or receive that which the parent gives him is one phase of the "filial" instinct.[7] The following question therefore arises. If a parent gives his child an object which the child adds to his collection would the child's behavior be due to the "collecting" instinct, or would it be due to the "filial" instinct, or would it be due to an acquired mode of reacting, or to several interacting urges?

As a *method* of "collecting," the term "Finding," (listed by Burk) suggests that much collecting activity has hitherto been classified as a manifestation of the "hunting" instinct. Moreover, when Miss Whitley lists "just take them," as a method of collecting the writers wonder whether this aspect of collecting may not be a twin brother of what many psychologists have been labelling the "fighting" instinct.

There is obviously a distinct need for a clearer and less ambiguous treatment of the behavior involved in "making collections" if this phrase is to have a status of respectability beyond that of a "faculty."

Investigation

The writers admit the limitations of the first method employed by them. The procedure was simply this: Two hundred numbered items (play activities) were placed in a check-list. The general directions for checking read as follows:

> What things have you been doing during the past week just because you wanted to? Read through the following list and . . . draw a circle . . . around each number that stands in front of anything . . . that you have been doing during the past week just because you wanted to do it.

Item No. 41 dealt with collecting activities. Cogent criticism of this prodecure as a means of studying children's interest in collecting will be found by reference to Miss Whitley's article.[1] One of the present writers has more recently studied "collecting" by means of a more specific technique. Much of the work has been carried on by students of the seminar in educational psychology at the University of Kansas. A list of one hundred fifty items was assembled. This list included *things** which children in the vicinity of Lawrence, Kansas, might find accessible and might therefore collect. Collecting was defined as an actual attempt on the part of the child to extend the

* By "thing" the writer meant any item that could be seen and carried away.

number and the kinds of items listed. Large numbers of children were then asked to check the list. Specific directions were given to the teachers as well as accurate ones to the children. The directions which were given the pupils follow:

Directions to Pupils

We are anxious to see what you have been collecting. Listen and follow directions carefully.

Section I.—Draw a ring around the number in front of every item you have collected at any time in the past, but which you are not now collecting. Example: Several years ago, Jack was collecting shells to see how many and how many different kinds he could collect. He is no longer hunting to find shells for his collections. If you have collected shells but are not now collecting shells, draw a line around No. 154. Do you all understand? Now draw a ring around the number of every item you have collected but are not now collecting. (Allow time.)

Section II.—Now go back over the list and *underline* those items which you are now actively collecting with the idea of seeing how many and how many kinds you can get together. (An example is here given and precaution is taken to insure understanding.)

Section III.—Write in the spaces at the end of the list any things you are now collecting which are not in the list.

Section IV.—Go over the list and put the letter S before everything you have collected *to sell.*

The list of one hundred fifty items was first tried out by Mr. Frank Boone upon several hundred children. The list was then lengthened to one hundred ninety items. These objects are ones, therefore, which children report that they collect. The writers do not wish to defend the children's definition of collecting or to discuss further the question of the instinctiveness of the tendency. It is obvious that the problem of the instinctiveness of the collecting tendency can hardly be settled so long as both terms ("instinctiveness" and "collecting") remain in the definition stage.

Table I.—Gross Data Regarding Children's Active Collections (After Boone)

	Boys	Girls	Both
Total number of individuals	808	868	1,676
Total number of collections	5,685	7,161	12,846
Average number per individual	7	8.2	7.6

Using the check-list of one hundred ninety items, Mr. Boone studied 1,676 children.[8] Most of the children were tested in March,

TABLE II.—LIST OF CHILDREN'S ACTIVE COLLECTIONS WITH NUMBER OF BOYS AND NUMBER OF GIRLS COLLECTING EACH ITEM (AFTER BOONE, 1927)

	808 boys	868 girls		808 boys	868 girls
1. Acorns	25	20	51. Compositions	11	38
2. Arrowheads	52	6	52. Crystals	17	19
3. Autographs	38	151	53. Cups and saucers (kinds of)	14	30
4. Autograph sentiments	22	121	54. Curios	34	28
5. Badges	50	18	55. Dishes (broken)	6	33
6. Bells	16	11	56. Dolls	5	99
7. Blank bank checks	30	20	57. Dolls of paper	6	126
8. Beads	22	105	58. Doll buggies	2	27
9. Beans (kinds of)	28	20	59. Doll clothes	3	101
10. Bees	14	11	60. Doll dishes	3	69
11. Beetles	22	9	61. Doll hats	2	51
12. Belts	23	17	62. Doll quilts	2	71
13. Birds' eggs	35	17	63. Dogs (toy)	6	14
14. Birds' beaks and claws	11	1	64. Dragonflies	10	12
15. Birds' nests	25	15	65. Drawings (number of kinds)	51	112
16. Birds' wings	14	5	66. Earrings	7	65
17. Blocks	17	13	67. Feathers	42	44
18. Bones	25	11	68. Ferns	12	37
19. Books, story	69	26	69. Fish hooks	73	15
20. Books, school (over years)	47	130	70. Flags	33	36
21. Books, song (kinds of)	25	76	71. Flint	48	10
22. Booklets	28	73	72. Flowers, pressed	27	100
23. Bottles	67	34	73. Flowers, paper	13	57
24. Bottle caps	42	25	74. Fossils	22	15
25. Boxes	51	53	75. Furs	56	20
26. Bracelets	6	59	76. Games	55	91
27. Brass	36	9	77. Glass (pieces of)	18	30
28. Breastpins	8	28	78. Glasses	8	14
29. Bullets or cartridges	75	4	79. Gopher skins	31	4
30. Butterflies	51	46	80. Grass (kinds of)	5	12
31. Buttons	34	40	81. Guns, toy	40	5
32. Buttons, pictures	16	9	82. Hair (locks of)	2	27
33. Buttons, cuff	16	9	83. Handkerchiefs (different)	35	131
34. Calendars	26	59	84. History scraps	40	54
35. Cards (election, merit)	35	79	85. Horse shoes	61	27
36. Cats (different kinds)	6	12	86. Insects	39	26
37. Chalk	26	46	87. Iron	52	10
38. Charm strings	5	26	88. Jackstones	13	78
39. China (painted)	15	38	89. Jewelry	26	117
40. Cigar papers	32	18	90. Jugs	22	7
41. Cigar holders	22	10	91. Keys	66	39
42. Cigar ribbons	24	19	92. Kites	51	14
43. Cigar tags	32	6	93. Knives	76	20
44. Cigar tins	32	8	94. Labels	17	12
45. Cigar stamps	49	16	95. Lace	8	78
46. Clocks	21	14	96. Lead	57	17
47. Clothes-pins	14	18	97. Leaves (kinds of)	26	81
48. Clover, four-leaf	49	113	98. Letters	27	124
49. Cocoons	28	21	99. Maps	40	41
50. Coins	136	90	100. Marbles	238	41

TABLE II.—*Continued*

	808 boys	868 girls		808 boys	868 girls
101. Marbles (agate)	165	26	149. Snails	16	5
102. Match box tops	20	17	150. Sayings (all descriptions)	14	22
103. Matches	20	5	151. Seals	13	24
104. Medals (kinds of)	37	15	152. Seeds (kinds of)	32	49
105. Minerals	13	13	153. Shells	50	49
106. Monograms	4	13	154. Shoes (old ones)	7	20
107. Moss (kinds of)	11	24	155. Shot	30	2
108. Nails	62	15	156. Skeletons	11	5
109. Napkins	13	59	157. Skins	28	2
110. Neckties (number of kinds)	35	10	158. Snake eggs	9	6
111. Newspapers	140	103	159. Soil (samples of)	6	6
112. Newspaper covers	9	12	160. Souvenirs	49	105
113. Newspaper scraps	21	61	161. Spiders	13	4
114. Needles	16	29	162. Spools	34	43
115. Nutshells (kinds of)	13	20	163. Spoons	9	34
116. Oak balls	12	11	164. Stamps	146	89
117. Oil cans	25	4	165. Stones (kinds of)	43	46
118. Padlocks	38	11	166. Strings	16	33
119. Paint (kinds of)	16	10	167. Tags, shipping	10	2
120. Paint charts	11	17	168. Tags, tin	12	3
121. Pants guards	10	2	169. Tags, tobacco	12	4
122. Paper, colored	24	67	170. Teeth	12	18
123. Paper, tissue	8	33	171. Thread	6	39
124. Pebbles	38	57	172. Tickets	36	42
125. Pencils	56	77	173. Time tables	24	19
126. Pictures	61	176	174. Tin	34	9
127. Pieces of calico	4	28	175. Tinfoil	12	15
128. Pieces of quilts	7	39	176. Tobacco sacks	36	11
129. Pins, common	14	27	177. Tops	67	14
130. Pins, hat	2	11	178. Tops (great variety)	25	35
131. Pins, scarf	2	10	179. Trade marks	19	5
132. Pins, stick	2	8	180. Trees, (samples of)	26	17
133. Poems (collected)	43	129	181. Valentines	81	242
134. Programs (dance, etc.)	19	69	182. Vases	19	33
135. Puzzles	44	70	183. Wagon wheels	41	9
136. Rabbit ears	48	7	184. Wampum	9	2
137. Rags	19	20	185. War relics	39	12
138. Rats, white	14		186. Water colors	28	46
139. Riddles	54	97	187. Wood (samples of)	26	19
140. Rings	22	89	188. Wood, petrified	30	20
141. Rocks (kinds of)	37	76	189. Wrappers, fruit	15	22
142. Rose petals	12	39	190. Wrappers, candy	44	44
143. Rubber	27	13	191. Milk bottles to be sold	2	4
144. Rubber sheets	6	2	192. Junk to be sold	4	
145. Sachet bags	1	28	193. Water from different lakes and oceans		1
146. Sacks (kinds of)	17	6	194. Rabbit feet	2	
147. Samples (tooth paste, perfume, etc.)	42	119	195. Old hats	3	
148. Sand, colored	9	5	196. Keys for rewards	1	

Items 191 to 196 inclusive, in Table II, were those found in the blanks at the end of lists checked.

1927. More than 12,000 active collections were reported. Table I presents data regarding the number of active collections and averages for the sexes. Table II presents the number of boys and the number of girls who reported that they were actively collecting each item included in the list.

It is of interest that most of the children report that they are actively making collections. Particularly noticeable is the fact that CA ten is the age at which the maximum number of active collections is reported. Mr. Boone found that the average number of collections reported by girls of this age was eleven, by boys ten. From ages ten to sixteen, there was a gradual decrease in variety of items collected, but *at every age level, the girls reported more collections than the boys.*

The writers wished to verify Mr. Boone's findings. They therefore employed exactly the same procedure in September, 1927, and repeated the experiment in April, 1928.

TABLE III.—PER CENTS OF CHILDREN OF VARIOUS AGES WHO REPORTED THAT THEY WERE ACTIVELY MAKING COLLECTIONS

	Boone study March, 1927		Present study September, 1927		Present study April, 1928	
	Boys	Girls	Boys	Girls	Boys	Girls
Number of cases	808	868	500	500	500	500
CA	Per cent					
9	88	84	87	84	83	84
10	85	85	85	85	85	90
11	84	90	84	90	84	86
12	82	84	82	85	82	85
13	70	83	75	86	75	86
14	76	88	74	86	78	83
15	77	73	74	74	78	76
16	60	60	60	60	68	64

In September, 1927, five hundred boys and five hundred girls were therefore studied intensively by one of the present writers. There were approximately equal numbers of individuals in each age group from nine to thirteen inclusive. An average of eighty-five per cent of the girls and eighty per cent of the boys reported that they were

collecting in September, 1927. The experiment was repeated in April, 1928. The same children were employed as subjects. The averages were approximately the same as in September, 1927; the average for the girls was eighty-four per cent, for boys seventy-nine per cent. A rather general tendency to collect is thus apparent. Table III shows the per cents of children of various age levels who reported to different investigators that they were actively making collections. Table IV presents similarly the number of active collections for different age levels. The ten articles most frequently reported by the writer's subjects of CA's nine, ten, and eleven are listed in Table V.

TABLE IV.—AVERAGE NUMBER OF ACTIVE COLLECTIONS REPORTED BY CHILDREN OF DIFFERENT AGE LEVELS

	Burk study, 1900		Whitley study, 1927		Boone study, Mar., 1927		Present study, Septemter and April, 1928[1]	
	Boys, ave.	Girls, ave.	Boys, md.	Girls, md.	Boys, ave.	Girls, ave.	Boys, ave.	Girls, ave.
Number of cases	607	607	2097	2349	808	868	500	500
Ages								
6	1.2	1.9						
7	2.1	2.3	3	2				
8	3.5	4	6	6				
9	3.9	4.	7	8	8	10	8	10
10	4.4	4.4	9	8	10	11	10	12
11	3.4	3.3	9	8	8	9	8	9
12	3.	3	9	9	7	7	8	7
13	3.5	3.4	8	8	5	7	5	7
14	3	3	7	7	5	6	5	6
15	2.7	3	6	7	5	5	5	5
16	2.1	2.8	5	6	4	4	4	4
17	2	2.5	5	5				
18	...	...	5	7				

[1] The last two columns of Table IV contain averages for the combined data of September, 1927, and April, 1928.

It is evident that the present study substantiates certain of the findings of Miss Whitley and those of Mr. Boone. Particularly noticeable is the fact that when the writer's check-list was employed the maximum number of collections was reported by children of CA ten.

TABLE V.—THE TEN ITEMS MOST FREQUENTLY COLLECTED BY BOYS AND GIRLS OF CA NINE, CA TEN, AND CA ELEVEN

100 boys of CA nine		100 girls of CA nine	
September, 1927	March, 1928	September, 1927	March, 1928
Cigar stamps	Marbles	Dolls	Valentines
Pictures	Agates	Doll clothes	Paper dolls
Newspapers	Newspapers	Paper dolls	Doll clothes
Nails	Stamps	Drawings	Dolls
Tools	Valentines	Jackstraws	Drawings
Kites	Nails	Beads	Jackstraws
Bullets	Coins	Clovers (four-leaf)	Beads
Marbles	Bottles	Coins	Clovers (four-leaf)
Knives	Tops	Poems	Pictures
Cigar bands	Bullets	Pictures	Poems
100 boys of CA ten		100 girls of CA ten	
September, 1927	March, 1928	September, 1927	March, 1928
Marbles	Samples	Letters	Valentines
Kites	Marbles	Story books	Paper dolls
Coins	Newspapers	Drawings	Samples
Newspapers	Agates	Pictures	Dolls
Tops	Stamps	Samples	Doll clothes
Tools	Knives	Handkerchiefs	Handkerchiefs
Bullets	Cigar bands	Poems	Pictures
Nails	Valentines	Riddles	Story books
Furs	Riddles	Flowers	Poems
Knives	Calendars	Curios	Leaves
100 boys of CA eleven		100 girls of CA eleven	
September, 1927	March, 1928	September, 1927	March, 1928
Marbles	Marbles	Letters	Valentines
Funny papers	Agates	Story books	Autographs
Newspapers	Newspapers	School books	Story books
Coins	Coins	Pictures	Pictures
Iron	Stamps	Beads	Feathers
School books	Valentines	Flowers (pressed)	Drawings
Pictures	School books	Handkerchiefs	Iron
Agates	Pictures	Earrings	Samples
Knives	Tops	Curios	Handkerchiefs
Story books	Wood (kinds of)	Drawings	Jackstones

Since children of a given CA show considerable variability in MA* the writers assembled the children into CA groups and studied them according to their MA's. Data were secured for two hundred children of eight years CA, two hundred of nine years CA, two hundred of ten CA, and two hundred of eleven CA. The children of MA ten were found in each CA group to report a larger average number of collections than the children of any other MA level.

Because CA ten (and MA ten) appeared to be the age at which the greatest number of collections was made, the writers extended their investigation. Accordingly, an entirely new group of one hundred boys of CA ten and a new group of one hundred girls of CA ten were selected for intensive study. They were normal subjects mentally (IQ 90–110, Stanford-Binet). The check-list of one hundred ninety items was given them in September, 1928. An average of twelve different collections was reported by the girls, eleven different collections by the boys. The objects collected were similar in rank to those reported in Table V. The averages correspond closely also to those reported in Table IV for a different group of children. It was noticeable here also that, although the average number of collections for all children was about eight, the average for CA ten was eleven. Children of CA ten in the vicinity of Lawrence, Kansas, thus appear to make a greater number of collections than do children of any other chronological or mental age.

One of the present writers later (October and November, 1928) studied two hundred additional children of CA ten. He first asked each child to write the names of the things he was collecting. The number of collections reported was four for the one hundred boys of CA. ten; five for the one hundred girls of CA ten. The conspicuous drop in the number of collections reported when the data were obtained by this procedure gives clear evidence that the findings of a given investigation are to a large extent a function of the technique employed. The children asked often: "Just what do you mean by a collection?" The following remarks are indicative of the confusion in the child's mind when he is asked to name the things he is collecting:

Do my school books count? I always keep them.†
I have several neckties. Are they a collection?†
I am getting different kinds of leaves for school. Are they a collection?†
What about suits?‡
Are dresses a collection?‡

* For example, the children of CA eight ranged from five to twelve in MA.
† As defined by the present questionnaire this would be a collection.
‡ As defined by the present questionnaire this would not be a collection.

The children were told that a collection consisted of any variety of things that they were assembling voluntarily in order to extend the kind or number. Even such prompting did not raise the average. When the check-list was placed before the children, the girls indicated that they were collecting twelve items; the boys eleven. The check-list acted as a memory aid and appeared to indicate that the children were considerably disturbed by the questions. Again and again such questions arose: "Is this a collection? I have several pieces of jewelry. I want more. Is that the same as collecting leaves?"*

If the question had been modified to this extent: Do you collect things just for the pleasure of having about you things you have acquired recently so that you can feel, taste, smell, hear, or see them?, one wonders what the answers would have been. Many of the boys indicated that they were collecting bottles or newspapers in order to sell them.* One girl indicated that she had five dresses and wanted more in order that she might appear pretty.†

Obviously, one is forced to the conclusion that the assumption of the instinctiveness of activities such as collecting is premature. Surely, one cannot count upon the motivation of collecting as a unique and easily recognized mode of behavior that may be listed as an instinct. A very great variety of behavior may be listed under this rubric. The purse of the junk-dealer seems moreover to be fully as much of a driving force as does any "instinct to collect." There may of course be certain cases of reaction where the response is predetermined, but the direct and indirect effect of past experiences cannot be denied. If we accept the hypothesis that every reaction leaves its effect upon the organism, we can hardly conceive of collecting as a purely unlearned behavior pattern.

Miss Whitley's admirable array of data lend corroboration to the writers' belief that results are a function of the technique employed in obtaining them. For example, Burk used no check-list at all; Miss Whitley employed a check-list of sixty items; the present findings were obtained by using a check-list of one hundred ninety items. The children in each instance reported a large number of collections when the data were obtained by means of a check-list. It is plausible also that the number of collections is a function of such other variables as CA, sex, locality, season, and the liberality with which one interprets the phrase "to make a collection."

* As defined by the present questionnaire this would be a collection.

† As defined by the present questionnaire this would not be a collection.

Pedagogical Implications

Educators and psychologists have stressed for many years the importance of utilizing the "collecting urge" or the "instinct to collect and hoard" in motivating school endeavor. In 1891, G. Stanley Hall called attention to the fact that the child's interest in collecting might be applied practically to the school situation.[9] Hall asserted that scores of boys who collect stamps or coins have very little knowledge of the geography or history necessary to give a rational interpretation to their collections. Therefore correlations of the two activities appeared desirable. Hall said further that, although many boys gathered blindly and mechanically large numbers of birds' eggs, they possessed little or no knowledge of the birds. He found also that many children were gathering the tails of birds or squirrels without knowledge of their characteristics or manner of mounting, or even the most common hunter's knowledge of the habits of the prey. Much hoarding seemed to be as wasteful as that of the miser who hoards in order to get ready to live. Hall's own account of how collecting might be made an effective basis of classroom procedure follows:

> There might easily be in most country towns, if not in each school building, small working collections, made largely by the children themselves, illustrating local geology, woods, plants, birds, and animals, mounted or put up in the most educational way, and, with a few pictured books, made the basis of general or class exercises. This, we believe, may at least be suggestive in solving the great question how more and better natural science may be taught in schools of the lower grades.[9]

Miss C. F. Burk suggested a few years later that teachers encourage their pupils to gather objects of nature, and that they be stimulated to relate how and where they found their specimens.[3] It will be valuable to examine critically the extent to which collecting and hoarding could be used today. When the population of the United States was small and when the people lived in relatively remote districts the advice of Hall appears to have been more appropriate than it would be in the year 1929. In a world grown somewhat anxious about conserving wild flowers and wild bird life, it is questionable practice to encourage children to gather "objects of nature." Nature lovers have for years protested against the fact that beautiful hillsides and dales are being wantonly despoiled by vandals who "collect" wild flowers, shrubs, fruit blossoms, and anything else that is at once attractive and portable. The beautiful objects soon wither and fade and the common good appears to be impaired by this baneful habit

of indiscriminate "collecting." Certainly, ruthless collecting should be discouraged.*

Since childhood is the time at which habits are most easily formed, it might today be better to urge children *not* to gather "nature's stores" but to choose with discrimination. By this means it might be possible to educate the public ultimately to the point where ruthless despoilation of nature's beauty would partially cease. It is apparent, of course, that it is the *type* of collection that one makes, and the *method* that one employs, that are of vast importance. It is obvious too that guidance is needed. The first step is to study *what things children collect in various communities.* The next step is to evaluate the practice and guide it. Children of CA ten will need, in the communities studied by the present writers, careful further study.

It will be noticed that G. Stanley Hall advocated *cooperative* collecting by the children. It is well known that much collecting activity is *individualistic* endeavor. It is probable that many children attempt to assemble a larger number of objects or specimens more rare or more beautiful than those of their neighbors. Since the child must learn how to live with others, it seems plausible that, instead of making individual competitive collections, the cooperative collection of legitimate and desirable objects would be the best use that could be made of the tendency to collect.†

It would seem then that certain of the children's collecting activities might well be discouraged or eliminated; for example, the collection of birds' eggs, the collection of matches, etc., (see Table II), certain of them might well be redirected; *i.e.*, the tendency to collect as an individual enterprise; others might well be encouraged and made the basis of educational growth. The reader will readily recognize that the above formulæ have been applied previously to the child's innumerable forms of mental and physical activity. No new formula is needed. So much for the pedagogical implications of the collecting interest! The psychological aspect is of a different nature. It is

*"Ruthless" includes not only the wild flower vandal; it includes also those who "collect" such souvenirs as Pullman towels, hotel silverware, and similar objects.

† The following newsnote is of interest in this connection: "Previous great expositions have consisted of competitive exhibits, alike in steel, textiles, education and everything else. The sponsors of the Chicago World's Fair (of 1933) announce that 'the old idea will be discarded; that the fair will exemplify the principle of 'strength in union,' by collective exhibits. The relationship of iron, steel, cement and the like to modern life will each be displayed in collective exhibits, tracing the progress of industry from infancy to the present time."[10]

clearly evident that, like numerous other problems in psychology, this one concerning children's interest in collecting is still in the definition stage.

References

1. Whitley, M. T.: Children's Interest in Collecting. *Journal of Educational Psychology*, 1929, Vol. XX, pp. 249–261.
2. Lehman, H. C. and Witty, P. A.: "The Psychology of Play Activities." New York: A. S. Barnes and Company, 1927. Pp. XVIII + 242.
3. Burk, C. F.: The Collecting Instinct. *Pedagogical Seminary*, 1900, Vol. VII, pp. 180, 207.
4. Thorndike, E. L.: Educational Psychology (Briefer Course). New York: Teachers College, Columbia University, 1915, p. 20.
5. Norsworthy, N. and Whitley, M. R.: "The Psychology of Childhood." New York: The Macmillan Co., 1922. p. 54.
6. Ogden, R. M.: "Psychology and Education." New York: Harcourt, Brace and Co., 1926, p. 59ff.
7. Woodworth, R. S.: "Psychology: A Study of Mental Life." New York: Henry Holt and Co., 1921. p. 150.
8. Boone, Frank: "A Study of Collecting among Children in Kansas. Unpublished Master's Thesis on file at Watkins Library. The University of Kansas Lawrence, Kansas, 1927.
9. Hall, G. S. (Note by the editor): *Pedagogical Seminary*, 1891, Vol. I, pp. 234–237.
10. The World's Fair at Chicago. Newsnote in *School and Society*, 1929, Vol. XXIX, p. 669.

AN IMPROVEMENT IN THE THEORETICAL BASIS OF FIVE POINT GRADING SYSTEMS BASED ON THE NORMAL PROBABILITY CURVE

WALTER CROSBY EELLS

Stanford University

One of the earliest and most widely used standard distributions of the school grades A, B, C, D, E was one in which it was recommended that they should be distributed symmetrically in the proportions 7-24-38-24-7 per cent. This was proposed by Cajori in 1914[1] and the mathematical basis elaborated in detail. It is clearly explained in Rugg's "Statistical Methods Applied to Education."[2] Figure 1 will make the basis clear. If the total range of the normal probability curve be taken as five sigma, only 1.24 per cent of the area under it will be found beyond the limits of such an assumed range. If this range be divided into five divisions of equal width, each one sigma wide, it is easily found from tables of the probability integral that the relative areas between such division points, and servicable approximations to them, are as follows:

Mark	Per Cent of Area	Per Cent Approximation
A	6.06	7
B	24.17	24
C	38.30	38
D	24.17	24
E	6.06	7
	98.76	100

On this simple basis of equal divisions along the base line this distribution of grades or similar ones have been recommended and widely used. Rugg states that

> For practical purposes he regards a five-fold distribution with range equal to $\pm 2.5\sigma$ as reasonably representative. This range has been taken by other workers in the field, notably by Ayres in the design of his "Scale for Measuring Spelling Ability." It should be pointed out that Ayres was the first to make this particular kind of practical use of the normal curve in education, in his handwriting scale as well as in his spelling scale.[3]

But is this system as nearly ideal as it at first appears to be? When once grades of A, B, C, etc. have been assigned, they are then

[1] See reference under first footnote Table I.

[2] Rugg, H. O.: "Statistical Methods Applied to Education." Houghton, Mifflin, Co., Boston, 1917, pp. 216–217.

[3] Rugg: *Op. cit.*, p. 219.

treated alike in any one group. In a true five point system there can be no distinction between a high B and a low B. However it is obvious from the unsymmetrical nature of the B section that the average of all B grades will not fall *at the center of this B interval*, but *somewhat to the left of it.* The same fact is true for the A grades. Only the C grades will have their mean at the center of their interval. The real unit that is used, perhaps unconsciously, is not the width of the class interval on the *X*-axis but the *distances between the means of successive grades.* Whenever grade points, quality credits, honor points, or similarly named measures are assigned on the basis of

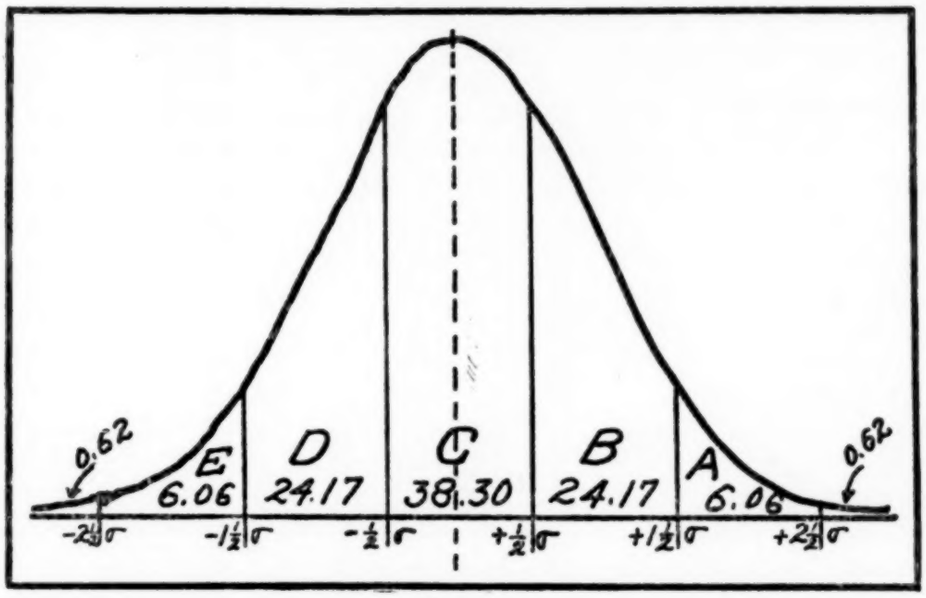


FIG. 1.

A = 3 grade points, B = 2 grade points, C = 1 grade point, this equality of successive intervals between means is tacitly assumed.

Kelley has given two formulas, one for the "mean deviation of the tail of a unit normal distribution" and the other for the "mean deviation of a portion of a unit normal distribution"[1] by means of which the actual distances in terms of sigmas from the center of the curve to the mean of any section of it can be computed when the per cent distribution is known. Applying his formulas to the suggested distribution of 7-24-38-24-7 per cents it is found that the distance from the center of the curve to the mean of all A grades is 1.9181 sigmas, to the mean of all B grades is 0.9105 sigmas. That is, the distance from an average A grade to an average B grade is

[1] Kelley, T. L.: "Statistical Method." Macmillan Co., New York, 1923. Formulas (53) and (55), pp. 100–101.

1.0076 sigmas, while the distance from an average B grade to an average C grade is only 0.9105 sigmas. Instead of equality of units along the *X*-axis, there is a ten per cent difference in the two distances! Should not a fundamental scientific measuring device have closer equality of units than this in different parts of the scale? It would hardly be considered scientific to have a yardstick vary from 36 inches to 39.6 inches.

Before presenting a system which will have the desired property of equality of fundamental units, it may be worth while to report

TABLE I.—VARIATIONS IN TERMS OF SIGMAS IN A-TO-B DISTANCES AND B-TO-C DISTANCES IN GRADING SYSTEMS PROPOSED BY VARIOUS INVESTIGATORS

Investigator and system proposed	Distance from mean of A's to mean of B's	Distance from mean of B's to mean of C's	Difference of two preceding distances
Cajori[1] 7-24-38-24-7	1.0076	0.9105	+0.0971
Dearborn[2] 2-23-50-23-2	1.2488	1.1711	+0.0777
Rugg[3] 3½-24-45-24-3½	1.1394	1.0683	+0.0711
George Washington University[4] 4-24-44-24-4	1.1109	1.0435	+0.0674
Cattell[5] 10-20-40-20-10	0.8940	0.8610	+0.0330
Meyer[6] 3-22-50-22-3	1.1329	1.1352	−0.0023
Mendenhall[7] 5-20-50-20-5	0.9895	1.0732	−0.0837
Smith[8] 10-15-50-15-10	0.8065	0.9485	−0.1420

[1] Cajori, F.: A New Marking System and Means of Measuring Mathematical Abilities. *School Science and Mathematics*, Vol. XIV, 1914, p. 283. Also in *Science*, n. s. Vol. XXXIX, p. 874, 1914.

[2] Dearborn, W. F.: School and University Grades. *University of Wisconsin Bulletin*, No. 368, Madison, Wisconsin, 1910, p. 17.

[3] Rugg, H. O.: "Statistical Methods Applied to Education." Houghton, Mifflin Co., Boston, 1917, pp. 216–219. Same method as outlined in detail in this paper, except with assumption of 6 sigma for length of base instead of 5 sigma.

[4] Ruediger, W. C., Henning, Geo. N., Wilbur, Wm. A., (Committee): Standardization of Courses and Grades. *Science*, n. s. 1914, Vol. XL, p. 643.

[5] Cattell, J. M.: Examinations, Grades, and Credits. *Popular Science Monthly*, 1905, Vol. LXVI, p. 371.

[6] Meyer, M.: The Grading of Students. *Science* n. s., 1908, Vol. XXVIII, pp. 243–250. See also Strong, E. K. Jr.: "Introductory Psychology for Teachers." (Revised), Baltimore, 1922, p. 197. Foster, W. T.: Scientific vs. Personal distribution of college credits. *Popular Science Monthly*, 1911, Vol. LXXVIII, p. 405.

[7] Brooks, Robert C.: Uniformity of Grading in Colleges and Universities. *School and Society*, 1915, Vol. I, p. 134.

[8] Smith, A. G.: A Rational College Marking System. *Journal of Educational Psychology*, 1911, Vol. II, p. 385.

the results of testing in the same way several other systems which have been proposed at different times by various writers. The present writer has found in the literature at least eight distinct symmetrical systems for college and university grading proposed for an ideal distribution of five grades on the basis of the normal curve. The proposer of each of these, the distance from an average of A grades to an average of B grades, the distance from an average of B grades to an average of C grades, and the difference of these two distances are given in Table I.

From Table I it is seen that the system so widely used, as proposed by Cajori, is next to the poorest of the eight tabulated, judging by the difference of the two distances, as reported in the last column of the table. These should be zero, or nearly so. The systems proposed by Meyer and Cattell show the least difference in length of units, Meyer's being particularly good from this standpoint with only a 0.2 per cent variation. But the units he uses are rather long ones, 1.13 sigmas, leading to a very small per cent of A grades which is hardly in accordance with most common educational practice. Cattell's, on the other hand, are rather short, being less than 0.9 sigma.

New System Proposed

The possibility of a system in which the distance from the mean of the A's to the mean of the B's is equal to the distance from the mean of the B's to the mean of the C's will now be investigated. If such a one can be found, it is obvious from symmetry that the distances between the means of A, B, C, D, and E grades will all be equal, a highly desirable condition as has already been pointed out. Consider Fig. 2. Let the distance from the mean of the C's to the mean of the B's equal X, and let the distance from the mean of the C's to the mean of the A's equal Y. Using Kelley's notation, let the boundaries of the grade intervals be indicated by Z_1, Z_2, Z_3, Z_4, (Z_4 being at infinity) and the proportions of the area of the entire curve to the right of successives Z's be indicated by q_1, q_2, q_3, q_4. Then from the Kelley-Wood Table[1]

$$q_1 = 0.500 \quad Z_1 = 0.398942$$
$$q_4 = 0.000 \quad Z_4 = 0.000000$$

[1] Kelley, T. L.: *Op. cit.*, pp. 373–385.

Z_2 and Z_3 are as yet unassigned and undetermined, as are the q's dependent upon them. Following Kelley's formula (55),

$$x = \frac{Z_2 - Z_3}{q_2 - q_3}$$

$$y = \frac{Z_3 - Z_4}{q_3 - q_4} = \frac{Z_3}{q_3}$$

$$y = 2x$$

These equations are a complete expression in mathematical form of the required conditions. Since the q's are functions of the Z's, this system of *three* simultaneous equations involves *four* independent variables, X, Y, Z_2, and Z_3. Therefore an infinite number of simul-

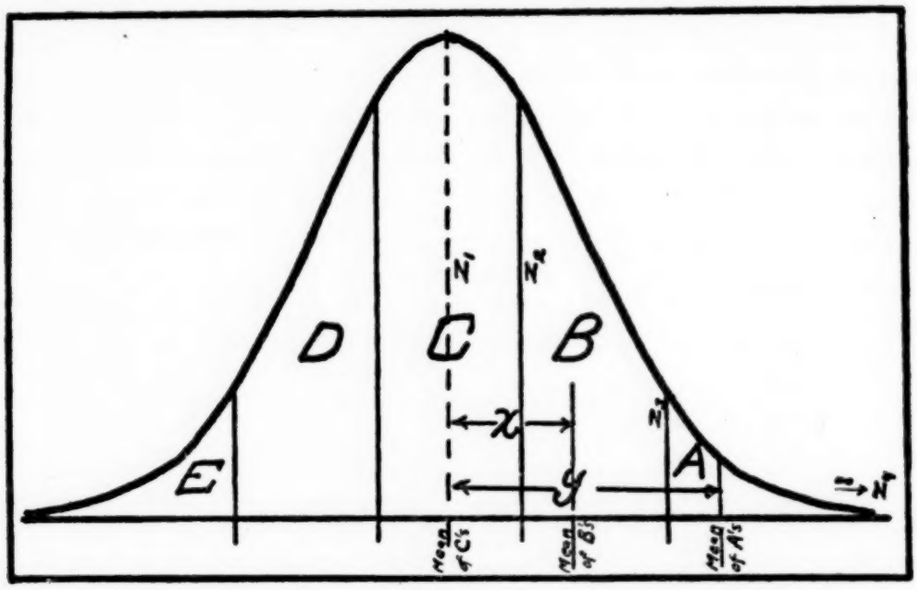


FIG. 2.

taneous solutions are possible; *i.e.*, not *one only*, but an *infinite number* of such ideal systems can be found, each having equality of distances between successive grade means. As soon as a value is assumed for X, the fundamental unit distance is determined, resulting in a unique solution of the system of equations given above.

As the simplest and probably most useful case, let $X = 1$ sigma. Then $Y = 2$ sigma, and $Z_3/q_3 = 2$. From column 5 of the Kelley-Wood Table it is found at once that this equation is satisfied (to the nearest 0.001) for $q_3 = 0.058$ and $Z_3 = 0.1160$, *i.e.*, when the per cent of A's is 5.8. Substituting these values of Z_3 and q_3 in the first equation, it becomes

$$X = \frac{Z_2 - Z_3}{q_2 - q_3} = \frac{Z_2 - 0.1160}{q_2 - 0.058} = 1$$

which reduced to

$$Z_2 = q_2 + 0.0580$$

This can be solved by trial, by successive approximations, with the aid of the third and fourth columns of the Kelley-Wood Table, securing as a solution, $q_2 = 0.277$ and $Z_2 = 0.3349$. Then the per cent of B's is 27.7 per cent − 5.8 per cent = 2.19 per cent and of C's is 2(50 per cent − 27.7 per cent) = 44.6 per cent. Thus a system in which the differences between successive mean values of the grades in the five intervals are each equal to one standard deviation, and a servicable approximation to the same which avoids fractional per cents, is as follows:

Grade	Exact, Per Cent	Approximate, Per Cent
A	5.8	6
B	21.9	22
C	44.6	44
D	21.9	22
E	5.8	6
	100.0	100

This is but one of the possible solutions of the given set of equations. Two others which might be preferred are secured by letting $X = 1.1$ sigma, and $X = 0.9$ sigma. Solving in the same way, but with some additional computation on account of the fractional sigmas taken for X, the following distributions are found which also have either exactly or approximately the desired property of equality of units on the X-axis between successive means of grades.

Grade	Fundamental unit 1.1 sigma, per cent		Fundamental unit 0.9 sigma, per cent	
	Exact	Approximate	Exact	Approximate
A	3.6	4	9.1	9
B	22.2	22	19.7	20
C	48.4	48	42.4	42
D	22.2	22	19.7	20
E	3.6	4	9.1	9
	100.0	100	100.0	100

Any desired number of systems can thus be constructed, each completely determined as soon as the length of the fundamental unit in terms of sigma is chosen. Since most of the systems recommended

and actually used throughout the country agree in recommending from five to ten per cent of A grades, the one in which the fundamental unit is taken as one sigma would seem to be preferable from the standpoints of simplicity, agreement with experience, and convenience of memory. *It is therefore recommended that the distribution of grades A, B, C, D, E be in the approximate proportions* 6-22-44-22-6 *per cents whenever it is desired to establish an ideal standard based on the assumption of normal distribution.* A further advantage over the Cajori system, and over others on the same basis, is that no part of the cases in the attenuated tails of the distribution is disregarded. While the A group is infinite in extent, the distance to the mean of all the individuals in it is finite and determined.

The value of the exact and approximate distributions recommended, as compared with those already discussed, is shown in Table II identical in form and interpretation with Table I.

TABLE II.—VARIATIONS IN TERMS OF SIGMAS IN A-TO-B DISTANCES AND B-TO-C DISTANCES IN GRADING SYSTEMS PROPOSED BY THE AUTHOR

System	Distance from mean of A's to mean of B's	Distance from mean of B's to mean of C's	Difference of two preceding distances
Eells:			
To nearest tenth per cent:			
5.8-21.9-44.6-21.9-5.8	1.0005	0.9994	+0.0011
3.6-22.2-48.4-22.2-2.6	1.0974	1.0991	−0.0017
9.1-19.7-42.4-19.7-9.1	0.8984	0.9008	−0.0024
To nearest per cent:			
6-22-44-22-6	0.9968	0.9886	−0.0082
4-22-48-22-4	1.0711	1.0827	−0.0116
9-20-42-20-9	0.9047	0.8996	−0.0051

It will be noticed from Table II that when per cents are carried out to tenths the largest discrepancy between the desired equality of A-to-B and B-to-C distances is about one-quarter of one per cent. When distributions are expressed in integral per cents only, as is more common and amply precise for all practical purposes, this discrepancy does not exceed one per cent appreciably in any one of the three systems proposed. In the most desirable one, where the unit distances are intended to be one sigma, the discrepancy is only four-fifths of one per cent. All six of the proposals outlined in Table II are superior to those examined earlier in this article, as summarized in Table I,

as regards the size of this discrepancy, with the single exception of the one proposed by Meyer.

SUMMARY

Seven different distributions of grades for a five-point grading system, proposed by different authors during the past quarter of a century, are shown to have discrepancies from four to sixteen per cent in relative length of units between successive grade means. A standard distribution of A, B, C, D, E grades of 6-22-44-22-6 per cents is proposed. If it were adopted as a standard, the unit distance between true means of successive grade groups, on the assumption of normality of distribution, would be uniformly one sigma, with an error of less than one per cent.

WEIGHTING RESPONSES IN TRUE-FALSE EXAMINATIONS

ELMER H. STAFFELBACH

San José State Teachers College, San José, California

The following is the report of a study to determine upon the most appropriate weights to assign to the three types of responses to statements in true-false examinations, namely, correct responses, incorrect responses, and omissions or statements not marked by the student.

Procedure of the Study

This study was carried out during the last half-semester of the school year 1926–1927. A true-false test consisting of 60 statements made up from the content of the seventh and low-eighth grade social science syllabus was submitted to 400 high-eighth grade pupils in social science. As a criterion of the pupil's true ability in social science, his average of seven half-semester marks in the subject was used. Only pupils who had on the school records seven half-semester marks in social science were included in the statistical portion of the study reported here. This eliminated 156 pupils whose records revealed less than seven half-semester marks, leaving 244 pupils to be included in the study.

As indicated above, the true-false examination was based upon the content which the pupils had been over during the seventh grade and low-eighth grade work in social science. The directions to the pupils included a statement that they might guess if they wished. They were informed, however, that they would be penalized for incorrect responses.

Raw Data Used in the Study

From the true-false test and the average of teachers' marks during seven half-semesters of work in social science, four types of scores were obtained for each pupil. These four types of scores are stated here with their representative symbols for the convenience of the reader. In each case X signifies a score. The subscript indicates the type of score.

X_0 = Average of seven half-semester grades in social science used as a criterion of the pupil's true ability in the subject.
X_1 = Number of statements responded to correctly in a true-false test taken from the content of the syllabus used during the time the criterion score was earned.
X_2 = Number of statements responded to incorrectly in the true-false examination.
X_3 = Number of statements not marked in the true-false examination.

Derived Data.—Further data necessary to the problem of weighting pupils' responses are given below:

r_{01} = Correlation between criterion and number of statements responded to correctly = .495.
r_{02} = Correlation of criterion with number of statements responded to incorrectly = −.602.
r_{03} = Correlation of criterion with number of statements not responded to = .061.
r_{12} = Correlation of number right with number wrong = −.642.
r_{13} = Correlation of number right with number omitted, −.458.
r_{23} = Correlation of number wrong with number omitted, −.104.
σ_0 = 10.32
σ_1 = 7.15
σ_2 = 6.80
σ_3 = 5.31

From these derived data, multiple regression equations of the following order were made:

$$\beta_{01.2} = \frac{r_{01} - r_{02}r_{12}}{1 - r^2_{12}}$$

These equations for the four-variable problem at hand were eight in number as follows:

$$\beta_{01.2} = .1847$$
$$\beta_{01.3} = .6616$$
$$\beta_{02.3} = -.6022$$
$$\beta_{21.3} = -.8725$$
$$\beta_{12.3} = -.6971$$
$$\beta_{03.2} = -.0016$$
$$\beta_{13.2} = -.5304$$
$$\beta_{21.2} = -.8927$$

From these results it was possible to compute regression coefficients of the following order:

$$\beta_{01.23} = \frac{\beta_{01.3} - \beta_{02.3}\beta_{21.3}}{1 - \beta_{12.3}\beta_{21.3}}$$

The three following regression coefficients necessary for our purposes were computed:

$$\beta_{01.23} = .3476$$
$$\beta_{02.13} = -.3617$$
$$\beta_{03.12} = .1831$$

From these results it was possible to compute the statistically most desirable weights by means of the following equations:[1]

$$W_1 = \beta_{01.23}\frac{\sigma_0}{\sigma_1} = .5017$$

$$W_2 = \beta_{02.13}\frac{\sigma_0}{\sigma_2} = -.5489$$

$$W_3 = \beta_{03.12}\frac{\sigma_0}{\sigma_3} = .3559$$

Meaning of the Weights

The significance of these weights is at once apparent. The highest possible correlation between the criterion and the test scores would be obtained by weighting the three types of responses as indicated above. Such weighting would, therefore, give the test its greatest value as a measure of social science ability, provided, of course, that our criterion be a valid one.

According to our results, if a correct response be given a value of .50, then an incorrect response would be given a negative value of approximately the same magnitude (−.549). An omission would be given a positive value of .356, approximately seven tenth as great as that given to a correct response.

Reduced to a weight of unity for a correct response, the weights would stand:

X_1 (correct response) = 1.00
X_2 (incorrect response) = 1.09
X_3 (omission) = .71

In other words in a test of three statements one of which is responded to in each of the three ways, the pupil's score would be:

$$+1.00 + .71 - 1.09 = .62$$

This weight of .62 in terms of percentage would of course become $100/3 \times .62 = 20 2/3$.

The Significance of the Results

It was not the writer's purpose in making the study to attempt to settle for all time the matter of weighting responses in true-false examinations; therefore, the results of the study are not to be interpreted in any such manner. However, since the conditions of the study were made with careful consideration to eliminate extraneous

[1] The statistical procedure used in this problem follows Kelley, Truman L: "Statistical Method." The Macmillan Company, 1923, p. 279ff.

factors, it is believed that for those conditions the weights arrived at above would be most nearly accurate.

In the final weights it will be seen that the incorrect responses would be weighted slightly more heavily (negatively) than the correct responses (positively). However, the difference is so slight as to be almost if not quite negligible. This bears out the usual chance theory of error in such tests, from which the right-wrong equation is derived; *i.e.*, $X = R - W$, or in terms of per cents:

$$X = \frac{100}{N}[N - (2W + O)]$$

when N = number of questions, W = number wrong, and O = number of omissions.

However, it will be seen that the weight derived for the statements not responded to is approximately seven-tenths as great as that for the correct response, and that it is likewise positive in nature.

This positive weight assigned to the omission tends to discredit the above per cent equation, since in the above equation an omission is given a weight of zero. The equation corrected for the weights determined in this study would be:

$$X = \frac{100}{N}[(R + \tfrac{7}{10}O) - W]$$

the symbols N, R, W, and O representing respectively the same as in the other equation.

Here, then, appears the tendency to reward the pupil who not only knows what he knows, but also knows what he knows not, when such happens to be the case. Considering the nature of the criterion, it may well be that this weight is due to a tendency on the part of teachers to rate highly the docile pupil who tends to take fewer chances scholastically. On the other hand, it may be assumed that our criterion of pupil knowledge in social science is a valid one, and that the weight is a reward for careful judgement or some other factor on the part of the pupil, such careful judgement or other factor operating in his regular school work to affect positively his gain in knowledge in social science.

A GRAPH FOR ESTIMATING RELIABILITY IN ONE RANGE, KNOWING IT IN ANOTHER

PHILLIP JUSTIN RULON

University of Minnesota[1]

The need for estimating coefficients of correlation and reliability for groups of varying dispersions has become sufficiently widespread to make simplified methods of such estimation a convenience to many.

Dr. Kelley[2] has supplied formulæ expressing the relationships which hold between coefficients of correlation and variabilities, in terms of the traits correlated, of groups in which the coefficients are calculated.

One such formula applies when different measures are correlated and when but one trait measured is curtailed.[3] This formula has been put into graphic form by Toops and Edgerton[4] and may be easily solved by the use of the abac, or computational graph, devised by these two writers.

Another formula[5] applies when similar or identical measures are correlated—and, accordingly, when both measures are curtailed. This formula applies, then, when dealing with reliability coefficients over differing ranges. The derivation of this formula may be found in the literature,[6] together with a discussion of its validity and the assumptions underlying its derivation.[7] The accompanying computational graph is based on this latter formula:[8]

$$\frac{\sigma}{\Sigma} = \frac{\sqrt{1 - R}}{\sqrt{1 - r}},$$

---

[1] College of Education and Institute of Child Welfare.

[2] Kelley, T. L.: "Statistical Method." The Macmillan Co., 1924, pp. 222, 225, and elsewhere.

[3] *Ibid.:* P. 225.

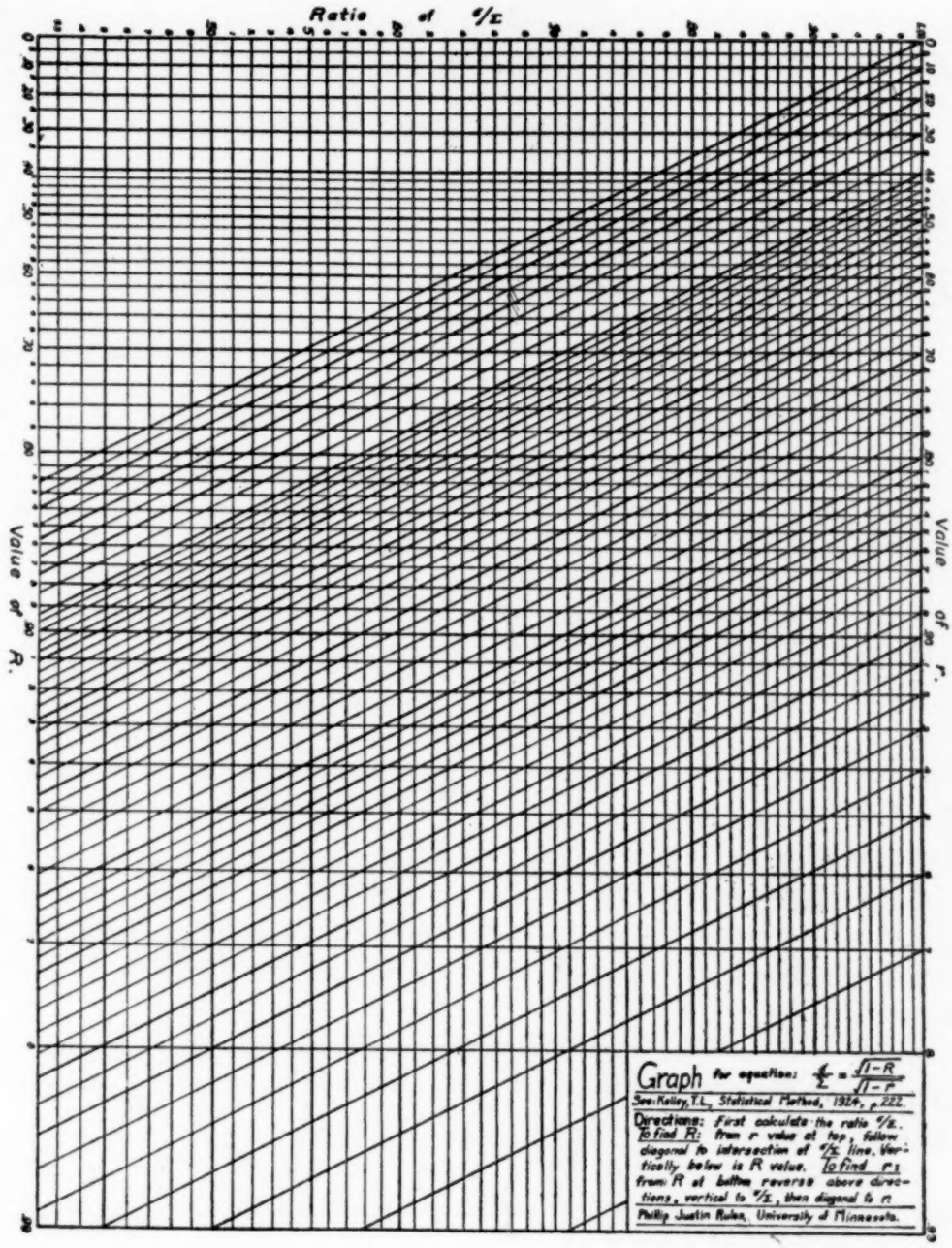
[4] H. A. Toops and H. A. Edgerton: An Abac for Determining the Probable Correlation over a Larger Range, Knowing it over a Shorter One. *Journal Educational Research*, Vol. XVI, No. 5, p. 382.

[5] Kelley: *Loc. cit.*, p. 222.

[6] Kelley, T. L.: The Reliability of Test Scores. *Journal of Educational Research*, Vol. III, No. 5, p. 370.

[7] Holtzinger, Karl J.: Communication, On the Assumption that Errors of Estimate are Equal in Narrow and Wide Ranges. *Journal of Educational Research*, Vol. IV, No. 3, p. 237.

[8] Taken from Kelley: "Statistical Method," (v.s.) p. 222.


Ratio of σ/Σ
Value of R.
Value of r.
Graph for equation: σ/Σ = √(1−R) / √(1−r)
See: Kelley, T.L., Statistical Method, 1924, p. 222.
Directions: First calculate the ratio σ/Σ. To find R: from r value at top, follow diagonal to intersection of σ/Σ line. Vertically below is R value. To find r: from R at bottom reverse above directions, vertical to σ/Σ, then diagonal to r.
Phillip Justin Rulon, University of Minnesota.

where:

σ is the standard deviation of one group in terms of the trait measured.

r is the coefficient of reliability in the range whose standard deviation is σ.

Σ is the standard deviation of another group in terms of the trait measured.

R is the coefficient of reliability in the range whose standard deviation is Σ.

This formula may be written:

$$\frac{\sigma^2}{\Sigma^2} = \frac{1 - R}{1 - r},$$

whence, taking logarithms,

$$\log \frac{\sigma^2}{\Sigma^2} = \log (1 - R) - \log (1 - r),$$

so that

$$\log (1 - R) = \log (1 - r) + \log \frac{\sigma^2}{\Sigma^2}$$

and

$$\log (1 - r) = \log (1 - R) - \log \frac{\sigma^2}{\Sigma^2}$$

These equations express a linear relationship, so that if abscissæ be laid off representing $\log (1 - R)$ or $\log (1 - r)$, then either of these may be found from the other by adding or subtracting a quantity determined by $\log \sigma^2/\Sigma^2$.

Such a relationship is represented in the accompanying graph, wherein the numerical values of r, R, and σ/Σ have been placed at the points representing the appropriate logarithms.

In using the graph, r and σ are always values applying to the smaller range, and R and Σ are those applying to the larger. If the reliability is to be estimated over a smaller range than the one over which it is already known, the graph should be used to find r; if the reverse holds, it should be used to find R. The ratio of σ/Σ is always to be calculated with the standard deviation in the smaller range as the numerator.

THE INFLUENCE OF WEIGHTED TRUE-FALSE TEST SCORES ON GRADES

JOHN GRAY PEATMAN

Columbia University

E. L. Clark recently[1] presented a formula by which scores on true-false tests could be transmuted so as to give *weighted scores* for each statement on a test and, consequently, a weighted total-score for each individual.

The formula: (Index of Validity) $IV = \frac{P - D}{1 - D}$ where D is equal to the per cent of subjects failing to answer correctly the particular true-false statement, and P is equal to the per cent of the "criterion group" missing the same statement. The criterion group consists of those subjects who are in the lowest "D per cent" of the class order-of-merit ranking, which order is determined by the usual unweighted grade on the test as a whole. For example, if forty of one hundred subjects taking the test miss statement number 1, $D_1 = 40$ per cent; of these forty subjects, if thirty are of the group comprising those forty subjects lowest on the entire test, then $P_1 = 75$ per cent; and IV_1 would be .58.

If the assumptions underlying the formula (see Clark's article) are true, then, for statement number 1, the number of subjects within the criterion group missing this statement is fifty-eight per cent as many as it could be in excess of the most likely number by chance.

Purpose and Procedure

The purpose of this study consisted in an attempt to ascertain the usefulness of this formula in obtaining weighted true-false test scores to be used in determining a subject's relative standing, or grade.

Through the cooperation of Professor R. S. Woodworth and Mr. R. G. Wendt, the author was able to obtain suitable data on Columbia University students enrolled in a graduate course in psychology. Data were obtained on six weekly quizzes of twenty-five true-false statements each and on the final examination of one hundred such statements. All of these examinations were given in successive weeks during the spring of 1929. Although the number of students taking each examination averaged about 100, the data for seventy-

[1] Clark, E. L.: A Method of Evaluating the Units of a Test. *Journal Educational Psychology*, Vol. XIX, 1928, pp. 263–265.

three students were used in the final computations because exactly this number took *all* of the tests.

In addition to the results obtained in attempting to arrive at an answer to the problem as stated above, the first two tables of statistical data are included primarily to make this study more complete for comparison with similar studies.

The procedure adopted:

1. The Index of Validity was obtained for each statement on the several tests.

2. The "value" of each test was determined in terms of the Mean of the Indices of Validity found for each set of statements.

3. The Mean (in terms of errors), Standard Deviation, and Pearson's Coefficient of Variation were calculated for the seventy-three subjects' original and weighted scores of each test, and also for the original and weighted *combined scores*, an individual's combined score being the average of his scores made on all of the examinations.

4. The subjects' original and weighted scores for each test and for the averages of all the tests (combined scores) were correlated by the product-moment formula, the data being treated in ungrouped arrays.

5. The degree of similarity between the two sets of *combined scores* was further investigated by subdividing the whole group of seventy-three students, arranged in order of merit, into various sub-groups, and finding the per cent of subjects identical to each sub-group when their order of merit under each set of scores was compared.

The results obtained:

1. *IV Value of Each Test.*—(If the assumptions underlying the development of the IV formula are true, then the IV value of a "perfect" test would be $1.00 \times N$ (statements), and thus the mean IV of all the statements would be 1.00. Professor Woodworth suggested that the "value" of each test be obtained in terms of the mean IV.)

IV Value of Each Examination

Examination	Number of statements	Mean	SD
A	25	.18	.18
B	25	.26	.16
C	25	.17	.13
D	25	.21	.19
E	25	.14	.18
F	25	.23	.18
Final	100	.21	.17

2. *Means (in Terms of Errors), Standard Deviations, and Coefficients of Variation for the Original and Weighted Scores of the Examinations.*

Examination	Original scores		Weighted scores		Original V	Weighted V
	Mean	SD	Mean	SD		
A	3.7	2.0	1.1	.64	52.7	56.0
B	3.3	2.8	1.0	.95	84.8	94.7
C	4.2	2.1	.8	.50	50.7	66.0
D	3.7	2.4	1.1	.78	63.6	71.2
E	4.7	1.2	.8	.49	26.5	63.6
F	4.0	2.5	1.0	.84	63.2	70.6
Final	20.8	8.5	5.7	2.61	40.8	46.0
Combined scores[1]	4.4	1.8	1.2	.53	39.8	45.9

(N students for each examination = 73. N statements for each = 25, except for final examination = 100.)

[1] In determining the averages for the combined scores, the final examination was given a weight of four, whereas each of the other examinations was given a weight of one.

3. *Correlation Coefficients (Original with Weighted Scores), Probable Errors, and Coefficients of Alienation.*

Examination	*r*	PE	*k*
A	.902	± .015	.43
B	.970	± .005	.24
C	.915	± .013	.40
D	.947	± .008	.32
E	.879	± .018	.48
F	.935	± .010	.35
Final	.955	± .007	.30
Combined scores	.978	± .003	.21

(N students = 73.)

4. *Per Cent of Similarity in Sub-groups.*—(In obtaining the following results, all seventy-three subjects were given an order of merit score (1 to 73) on the basis of their original combined-scores and also on their weighted combined-scores. They were then divided into two sets of sub-groups—one set based on original scores, and the other on weighted scores—and the per cent of the same subjects common to corresponding sub-groups was determined.)

(*a*) A distribution of sub-groups arranged in size according to the "normal" frequency curve:

Sub-group	Number of subjects	Per cent of similarity
A	3	100
B+	5	60
B	12	58
B−	16	88
C+	17	82
C	12	100
C−	5	80
F	3	67
		Average = 79

(*b*) A distribution of 7 sub-groups, 10 in a group:

Sub-group	Per cent of similarity
1–10	90[1]
11–20	80
21–30	90
31–40	70
41–50	70
51–60	90
61–73 (N = 13)	92
	Average = 83

[1] Which indicates that of the ten individuals ranking highest on the basis of their original scores, nine of these ten also ranked highest on the basis of their weighted scores.

Interpretation and Conclusion

1. Although all of the correlation coefficients are relatively high and all but one (Examination E) adequate for predictive purposes, the two coefficients obtained for the Final Examination and for the Combined Scores, being .955 ± .007 and .978 ± .003 respectively, are most important in indicating nearly perfect correlation between the original and weighted scores.

The coefficient obtained on the Final Examination scores is considerably more important than any of those obtained on any of the various quizzes because of the greater reliability accruing to the final examination of one hundred statements.

The coefficient obtained on the Combined Scores is even more important than the Final Examination coefficient because grades and rankings of students are determined on the basis of averages such as were used in getting this correlation of .978 ± .003 on the Combined Scores.

2. The per cents of similarity obtained when the subjects were divided into sub-groups reveal the degree of relationship between the two sets of scores more exactly than does the correlation coefficient of .978 ± .003, although the relationship given by the coefficient becomes somewhat more obvious in view of the Coefficient of Alienation, which was .21 in this case.

Dividing the subjects into sub-groups on the basis of their original scores, arranged in order of merit, and making the size of the groups correspond to the "normal" frequency distribution, seventy-nine per cent of the subjects, on the average, were found to be in corresponding sub-groups determined by their weighted scores. The similarity averaged eighty-three per cent when the comparison was made with seven sub-groups, ten subjects in a group.

Even when considered from the point of view of these small sub-groups, the degree of correspondence between the original and the weighted scores is very high.

3. On the basis of these data and the results obtained in the treatment thereof, the author tends to the conclusion that the weighting of these true-false scores by the method herein used gave results such as to indicate that the use of the IV formula *in determining these subjects' relative standing, or grades,* is not justified.

Its use is unwarranted not only because of the high correspondence between the original and the weighted scores, but also in view of the great amount of labor involved in its application. Although some of the statistical work is eliminated by the use of a table of IV's for the most common D's and P's (which Clark suggests and the author used), the relative amount of labor saved is not very significant. Furthermore, if the size of the group varies, certain saving devices are eliminated.

It should be added that this conclusion is not to be interpreted as vitiating the IV formula when used to ascertain the relative merits of a particular unit, *i.e.*, a true-false statement, used in a quizz. As stated earlier, this study was concerned with the usefulness of the formula in obtaining weighted true-false test scores to be used in determining a subject's relative standing, or grade.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY FRANCES M. FOSTER

The Child in America; Behavior Problems and Programs, by William I. Thomas and Dorothy Thomas. New York: Alfred A. Knopf, 1928. Pp. XI + 583.

Interest in child adjustment problems and facilities for the study of behavior deviates have grown so rapidly in recent years that a comprehensive treatment of the subject is gladly welcomed. All who deal in any capacity with the problem child will find here a wealth of detailed information about procedures and methods, and a critical evaluation of present practice. The authors deal with their subject from all possible viewpoints. The setting is prepared for the reader by the presentation in the first chapter of pictures of various types of maladjustment, largely in the form of case material. The authors indicate two primary causative factors in maladjustment—organic defect or abnormality in the individual, and wrong habit formation. The remainder of the book is devoted to methods of study and treatment of behavior deviates as indicated by present practice. Methods of dealing with delinquency through the court and the reform school are critically examined. An extensive account is given of psychiatric child guidance clinics and community organizations, club and recreational facilities, for dealing with child behavior problems. The authors are skeptical of what they term the psychometric approach in the study of problem children though, in their opinion, tests of special abilities and varied aspects of personality may be helpful. The psychiatric approach in its various interpretations is, in the author's opinion, an effective medium for the study and correction of behavior problems; the physiological approach gives much promise for the future; and the sociological approach, the main objective of which is the study of the individual in relation to others, holds promising possibilities as a diagnostic, correctional, and preventive measure.

The reviewer would commend the book as a sincere task admirably executed. It is too extensive for quick and easy reading, though justifiably so in view of the importance of the topic and the necessarily

large mass of facts presented. It is thought-provoking with little attempt on the part of the authors to settle matters arbitrarily through superficial recommendations. There is an attempt to present the facts without bias or predilection for particular methods or systems. Appreciation and constructive criticism of the contributions from all sources is freely expressed.

Although for convenience of presentation it may be necessary to make a distinction between the psychiatric, the psychometric, and the sociological approaches in child study, these divisions nevertheless set up artificial distinctions which are not justified by child behavior. The authors contribute to the reader's perplexity by treating psychometric and personality study methods as two different things, and the study of the child's behavior toward his comrades in the nursery school as quite another. This seems "far-fetched" to any one who is trying to understand the whole child.

The "IQ" is weighed in the balance and found wanting. The author's treatment of the subject undervalues standardized measurement as a scientific method of child study and observation. A discussion of psychological measurements with no reference to their function as fact-finding and diagnostic instruments in the hands of experts seems futile. The question of validity in connection with methods of personality study and the physio-morphological approach is scarcely raised. The authors have almost entirely overlooked child study activities of successful educational clinics, psychological clinics and vocational bureaus established in the larger school systems, universities, and community centers. Names of some of the foremost investigators of child study problems in these organizations do not appear.

In spite of these shortcomings, the book well repays the reader and will prove to be a rich source of reference material.

GERTRUDE HILDRETH.

The Lincoln School of Teachers College

The Commonwealth Teacher-training Study, by W. W. Charters and Douglas Waples, with an Introduction by Samuel P. Capen. Chicago: University of Chicago Press, 1929. Pp. XX + 666.

During the last decade at least two-score books have been published which deal directly with curriculum problems. In so far as periods of time may be described in terms of educational movements, we may readily call the past ten years a "curriculum decade."

Outstanding among those who have fostered this movement is Dr. W. W. Charters, who has contributed largely to the development of curriculum-making procedures and who has applied those procedures in the solution of numbers of practical curriculum problems. Among these we may mention his work in the development of curricula in retail selling, secretarial work, librarianship, and in the teaching of ideals.

This scientific movement in the field of curriculum making has taken a tremendous step forward this year, with the publication of "The Commonwealth Teacher-training Study."

There are two reasons for the unusual significance of this study. In the first place, it is of such a size, scope, and scientific character as to become a model for future curriculum research, in whatever field or subject. In the second place, it is the first and only comprehensive effort to collect basic data regarding curricula for teacher training.

The investigation upon which the present book is based was made possible by grants from the Commonwealth Fund and by the cooperation of the University of Chicago, and extended over a period of three years. Cooperation was secured from educational leaders throughout the nation, and on a large scale. It is therefore sufficiently comprehensive and representative of the country as a whole to be of nationwide service.

Briefly, the authors have done three things: (1) They have compiled a complete master-list of teachers' traits. (2) They have compiled a similar master-list of teachers' activities. (3) They have obtained statistical evaluations of the several items in the lists so that we may know the relative amounts of emphasis to place on each in the training of teachers.

The portion of the study which deals with teachers' traits is quite brief, and has resulted in a final grouping of all the traits under twenty-five main heads, such as adaptability, attractiveness, breadth of interest, carefulness, considerateness, cooperation, etc. These have been ranked by expert judges separately for teachers in different school grades and situations.

The greater bulk of the published report deals with the activities of teachers. This is unquestionably the most complete activity analysis of teachers that has ever been made. Lists were collected from all available previous studies and from several thousand individual teachers. These were classified into divisions, subdivisions,

sections, subsections, etc., until the analysis reached the degree of detail judged to be best suited for instructional purposes.

Having prepared a classified master-list of teachers' activities, the investigators evaluated the items by having each one rated as to frequency, difficulty, importance, and need of being taught in the teacher-training school. Such ratings were secured with respect to different grade levels, thus yielding different rankings of the items for kindergarten, rural, high-school, and other teachers. The authors are to be commended for their use of exact and rigid statistical methods of determining the reliability of the ratings they have secured.

The last one hundred sixty-two pages are devoted to summary tables showing the values of the various items. To the practical curriculum maker, these tables will be the most useful parts of the book. They present the items classified to show how they are judged by different types of experts, how they apply to different kinds of teachers, and also in such a way as to distinguish between frequency, difficulty, importance, need of teaching in school, and a composite of all these.

In addition to the strictly objective material already described, there is a goodly amount of sound curriculum theory, with very definite and helpful suggestions as to how the data may be used in revising existing courses and in constructing new ones. When we add to these possibilities a third one, namely the possibility that this same procedure may be applied in other fields than teacher-training, it is apparent that "The Commonwealth Teacher-training Study" is the outstanding event of a decade of curriculum work.

C. C. Crawford.

University of Southern California.

Race Attitudes in Children, by Bruno Lasker. New York: Henry Holt and Co., 1929. Pp. XVI + 394.

"Race Attitudes in Children" is so well organized, the problems so plainly stated, and the conclusions so clearly drawn, that the outline can be easily presented. The study was made "under the direction of Bruno Lasker by members of The Inquiry, a national organization for the promotion of cooperative studies of problems of human relations."

In his Preface, the author summarizes the contents of the book as follows: "The first part shows what the race attitudes of children are,

the second how they have been acquired, the third what intentional teaching has gone into their formation, and the fourth how they may be modified." As a result of the study, Lasker's outstanding conclusions relative to the first point are that, with respect to small children, fear, cruelty, combativeness, and ridicule are the predominant attitudes toward members of certain other races. The author finds the types of attitudes noticeable among adolescents somewhat different, and lists them as condescension, rivalry, class consciousness, and returning fears.

As to the source of racial attitudes, Lasker minimizes the factor of instinct, taking the position that they are, in the main at least, learned behaviors. He replies to the question "What intentional teaching has gone into their formation?" with the statement that much of the teaching in our public schools is largely propaganda, and that public speaking or oral English is "particularly nefarious" in this respect. And, of course, he has a word for those parents and Sunday School teachers who inculcate racial prejudices in children. Finally, in the matter of the very difficult task of modifying racial prejudices, Mr. Lasker suggests that it can be brought about in part, at least, through teaching, but that the emphasis must be placed upon *method* rather than *content*.

The reviewer has only praise for "Race Attitudes in Children." A very difficult and intangible subject has been handled clearly and honestly. To be sure, the evidence that is advanced in the book is far from being scientific, but the author frankly states that he makes no claims for it on that basis. Mr. Lasker has treated his material in a fashion that should give pause to those adults who are cluttering up the minds of children with attitudes of antagonism and prejudice toward those of other races.

HERBERT A. CARROLL.

Teachers College, Columbia University.

Educational Psychology, an Objective Study, by Peter Sandiford. London: Longmans, Green and Co., 1929. Pp. XIX + 406.

This is an educational psychology written from the viewpoint of behaviorism. It is the more valuable in that it is not forever trumpeting this fact. The author cannot restrain one reference to metaphysicians who revel in introspective orgies, but for the most part he proceeds calmly enough with his task of presenting the available objective data.

The book is divided into Part I, "Man's Equipment for Learning," and Part II, "The Learning Process." Part I begins with an excellent chapter on heredity. The author then feels that he must describe how the human body works, and we have chapters on the effectors, the sense organs, and the nervous system. While this material may serve to make the book a treatise complete in itself, and may make it more usable in normal schools where educational psychology is the only psychology taught, the reviewer feels that educational psychology is properly an advanced course. And these chapters will add to that confusion and overlapping of courses which is a scandal in departments of psychology today. After the student has taken courses in general, social, and educational psychology and has studied the nervous system three times, he ought to know it fairly well! There follow chapters on instincts, emotions, and intelligence, the last proving that it is possible for a behaviorist to present a just and adequate treatment of the testing movement.

Part II, "The Learning Process," begins with a chapter on the conditioned reflex. The material on learning, retention, fatigue, and transfer (which some writers can make so frightfully dull!) is skillfully and interestingly summarized, and the reader is not disappointed in his expectation of concrete objective facts. There is a minimum of the theorizing that infests so many texts, behavioristic and otherwise. The book concludes with a treatment of educational tests and summaries of investigations in reading, spelling, handwriting, and arithmetic. The reviewer was disappointed not to find similar treatments of other subjects.

The writer is a professor in the University of Toronto, and the book, while based in the main upon research in the United States, possesses a trace of foreign flavor which adds to its attractiveness.

Teachers will find this book eminently usable as a text. Its behavioristic bias is a clear gain in as much as there has resulted an admirable program of sticking to the facts. The subtitle, "An Objective Study," is no empty boast. At the same time the book is so little doctrinaire that it will offend nobody. Indeed, the author's treatment of emotions and of the laws of learning is somewhat confusing from his failure to reduce the material to a clear cut schematism of his own. But this may be more valuable than a too facile resolution of difficulties. The reviewer believes the book will find extensive use.

MELVIN RIGG.

Kenyon College.

Creative Power, by Hughes Mearns. New York: Doubleday, Doran and Co., 1929. Pp. 396.

Bismillah! there is only one Hughes Mearns and Mearns is his prophet. Mearns is still his prophet. True, there are any number of enthusiasts, not to say zealots, charging about by this time in the desert of creative writing for children. But the word still lies with Mearns; when he opens his lips, the impossible becomes real, the improbable becomes logical. Children trust him, he releases their powers, he unlocks their expression, he lures them into confession, he coaxes them into communication.

His secret is simple. It lies, not in his clear mind, though his is of the clearest, nor in his plausible organization of method. Any of your college-entrance grammarians might rank him in clarity and administrative knack. Hughes Mearns has somewhere got hold of the only magic that will serve the—what shall I call 'em?—the Leopold Auers of the arts. Somehow, by force of "personality," by his almost Irish voice, by—by whatever it is, he manages to persuade youngsters of the reality, of the importance of their private—even of their secret—yes, even of their ridiculous experiences. And then—well, then they've got something which simply must be said, and they proceed to say it.

He's really your only possible kind of teacher in the arts—the artist! He will deny that he is a "teacher." Indeed, he does deny it:

"I am a writing man and not a teacher of the subject of writing; and there is a vast difference between these two sorts of creature."

He stands Bernard Shaw's flippant epigram on its head; he not only can and does, he can and teaches, too! More power to him, say all the rest of us. And therefore the more regretfully do all the rest of us see the risk he runs cavorting about the country when he should be home in a magic hill working direct magic. If æsthetic power, like religious vocation, comes from the immediate laying on of hands, Mr. Mearns should be dealing at first hand with the potential creators, with the children themselves. He should be teaching children; or, at worst, he should be further sensitizing some few choice spirits like himself who already have some measure of his own gift. His mass-conversions à la Charlemagne can result only in pedagogic distortions, in zealous thumbhandedness throughout English Departments—God wot!—with "creative" 20-minutes planned for in cold curricula and courses of study, after the best, well-meaning, unimaginative manner of the conscientious system-

administrator. And Mearns himself in that event? Alas, Cassandra; Mearns will lose touch with the very source of his own magic; he will lose touch with the children; he will codify "principles," and "objectives," and "laws," and all manner of sound, logical nonsense; and those will all be learned and easy to learn, and all based on blasted reminiscence, if you please, and so not worth the paper to print 'em on.

When "Creative Youth" appeared in 1925, its impact on English method was instant and startling. It was A Book—by now it is almost a Bible. One hesitates therefore to bay even more incoherently about "Creative Power." Yet "Creative Power" is much the better book of the two. After all, the scope of "Creative Youth" comprised the peculiar and therefore narrow field of poetic expression by children. "Creative Youth" showed what children can do in the lyric, the rhythmic, if you like. "Creative Power," on the other hand, while apparently handling, even repeating the same problem, is the work of a ripened and matured Mearns, based on a much wider, much wiser philosophy. The new book is really an unconventional, highly amusing treatise on education itself, in its broadest significance; readable, sound, human, and quotable almost in entirety. There is so much common-sense in it, for instance, on the vital, irritating question of the place of the youngster in the group, that the book emerges, because of this treatment alone, from the professional classification of handbooks on English Composition, and takes quite a decent place among more pretentious-seeming tomes on life and living. Why under the shining canopy Mr. Mearns christened it "Creative Power," therefore, must be a secret between him and the blurb department of Doubleday Doran. A fine, even a *great* book—a rotten title. Perhaps the ghost had to be walking. Rest, rest, perturbed Babbit.

B. J. R. Stolper.

The Lincoln School of Teachers College.

PUBLICATIONS RECEIVED TO JANUARY 15, 1930

Monographs

Baldwin, J. W.: *The Social Studies Laboratory.* New York: Bureau of Publications, Teachers College, Columbia University, 1929, pp. 98.

Fräßdorf, Walter: *Die Psychologischen Anschauungen J. J. Rousseaus.* Langensalza: Hermann Beyer und Söhne, 1929, pp. 248.

Offner, Max: *Die Geistige Ermüdung.* Berlin: Verlag von Reuther und Reichard, 1928, pp. 180.

Sauvage-Jousse, Paul: *Le Métaverbe.* Paris: Librairie Felix Alcan, 1928, pp. 222.

Van Alstyne, Dorothy: *The Environment of Three-year-old* Children. New York: Bureau of Publications, Teachers College, Columbia University, 1929, pp. 108.

Welkisch, Artur: *Exper.-vergl. Untersuchungen über die sittliche Bildung in d. verschiedenen Volksschulformen.* Langensalza: Hermann Beyer und Söhne, 1929, pp. 92.

Textbooks

Clark, John R., Otis, Arthur S., and Hatton, Caroline: *Modern-school Arithmetic,* Book One. Yonkers: World Book Co., 1929, pp. 399.

Clark, John R., Otis, Arthur S., and Hatton, Caroline: *Modern-school Arithmetic,* Book Two. Yonkers: World Book Co., 1929, pp. 386.

Clark, John R., Otis, Arthur S., and Hatton, Caroline: *First Steps in Teaching Number.* Yonkers: World Book Co., 1929, pp. 225.

Cochran, Grace, and Eddy, Helen M.: *Si Nous Lisions.* Chicago: University of Chicago Press, 1929, pp. 192.

Eddy, Helen M.: *Beginning French.* Chicago: University of Chicago Press, 1929, pp. 275. French Workbook to Accompany "Beginning French," pp. 155.

Passarelli, Luigi A., and Pezard, André: *Les Drôles Aventures de Renard.* Chicago: University of Chicago Press, 1929, pp. 141.

Theisen, W. W., and Leonard, Sterling A.: *Real Life Stories, a Seventh Reader,* Volume VII. New York: The Macmillan Co., 1929, pp. 487.

THEISEN, W. W., and LEONARD, STERLING A.: *Real Life Stories, an Eighth Reader,* Volume VIII. New York: The Macmillan Co., 1929, pp. 504.

ADMINISTRATION

CUBBERLEY, ELLWOOD P.: *Public School Administration.* Revised and Enlarged Edition. New York: Houghton Mifflin Co., 1929, pp. 710.

KNIGHT, EDGAR W.: *Education in the United States.* Boston: Ginn and Co., 1929, pp. 588.

MCGREGOR, A. LAURA: *The Junior High School Teachers.* Garden City, N. Y.: Doubleday, Doran and Co., 1929, pp. 284.

REEVES, FLOYD W., and RUSSELL, JOHN DALE: *College Organization and Administration.* Indianapolis: Board of Education, Disciples of Christ, 1929, pp. 324.

SMITH, HARRY P.: *Business Administration of Public Schools.* Yonkers: World Book Co., 1929, pp. 432.

TROXEL, OLIVER LEONARD: *State Control of Secondary Education.* Baltimore: Warwick and York, Inc., 1928, pp. 232.

MISCELLANEOUS

BROWN, MAUD A.: *Teaching Health in Fargo.* New York: The Commonwealth Fund, Division of Publications, 1929, pp. 142.

BURTT, HAROLD: *Psychology and Industrial Efficiency.* New York: D. Appleton and Co., 1929, pp. 395.

FLORENCE, P. SARGANT: *Sociology and Sin.* New York: W. W. Norton and Co., 1929, pp. 98.

HATFIELD, H. STAFFORD: *The Conquest of Thought by Invention.* New York: W. W. Norton and Co., 1929, pp. 80.

JEFFREYS, H.: *The Future of the Earth.* New York: W. W. Norton and Co., 1929, pp. 79.

NELSON, LOUISE A.: *Variations in Developments and Motor Control in Goiterous and Non-goiterous Adolescent Girls.* Baltimore: Warwick and York, Inc., 1929, pp. 193.

O'SHEA, M. V.: *Newer Ways with Children.* New York: Greenberg, Publisher, Inc., 1929, pp. 419.

SAUDEK, ROGERT: *Experiments with Handwriting.* New York: William Morrow and Co., 1929, pp. 395 + Plates.

WALKER, K. M. and E. M.: *On Being a Father.* New York: W. W. Norton and Co., Inc., 1929, pp. 192.

FRANZEN, RAYMOND: *Physical Measures of Growth and Nutrition.* New York: American Child Health Association, 1929, pp. 138.

TESTS AND EXERCISES

BLAISDELL, J. G.: *Instructional Tests in Biology for Junior and Senior High School.* Yonkers: World Book Co., 1929, pp. 56.

CARMAN, H. J., BARROWS, THOMAS N., and WOOD, BEN D.: *Junior American History Test.* Yonkers: World Book Co., 1929, Forms A and B.

CHEYDLEUR, FREDERIC D.: *The Relative Reliability of the Old and New Type Modern Language Examinations.* Reprint from "The French Review," May 1929, Vol. II, No. 6, pp. 22.

CLARK, JOHN R., OTIS, ARTHUR S., and HATTON, CAROLINE: *Modern-school Individual Number Cards.* Yonkers: World Book Co., 1929.

LEONHARDY, ALMA, HOGOBOOM, GRACE W., and VAN PATTEN, ELIZABETH: *Directed Study Guides for Pitkin and Hughes' Seeing America Farm and Field.* New York: The Macmillan Co., 1929, pp. 64.